schwa

Linguistic Perspectives on Language

Schwa: Linguistic Perspectives on Language is published by the Department of Linguistics and English Language at Brigham Young University. Its contents represent the opinions of the authors and editors, and not necessarily those of the faculty or administration of Brigham Young University.

We are always looking for staff members and articles. If you are interested, please e-mail us at schwa.journal@gmail.com or check out our new website at schwa.byu.edu.

Table of Contents

Students

Staff

Editor in Chief
Stetson Robinson

Managing Editors
Aaron Airmet
Lauren Bangerter

Senior Editors
Carli Hanson
Dana Knudsen
Greg Scoggin
Heidi Bishop
Jessica Allen

Art Director
Christina Johnson

Senior Designer
Whitney Hanks

Editors
Bethanie Thatcher
Brooke Weaver
Bryn Clegg
Christina Rosenquist
Diane Cardon
Jackson Tetmeyer
Jennifer Johnson
Jill Bickham
Lauren Atkinson
Mark Melville
Matt Garrett
Rachel Brown
Sara Guggisberg
Sara Townsend
Scott Mosley
Shamae Budd
Tiana Cole
Victoria Candland
Whitney Hanks

Designers
Jill Bickham
Sara Guggisberg
Victoria Candland

Editor's Note

Three cheers for the senior linguitics student symposium!

Near the end of each semester, senior linguistics students in the Senior Seminar course (Ling 490) present their well-crafted theses at a student symposium held on campus where anyone may attend. The topics at the symposium vary widely along a spectrum of compelling linguistic issues in phonetics, syntax, literacy, semantics, processing, psycholinguistics, language acquisition, and more. Attenders (or readers) are bound to find something in their field of interest.

Furthermore, starting with this issue, *Schwa* now features a section of top-pick theses presented at the senior linguistics symposium from the previous semester. In this case, our highlighted essays include Coleman Bonham's "Could Watson Pass the Turing Test?," Michael Wyatt's "Linguistic Analysis of . . . Black Vernacular English in *The Help*," and several others we hope you will enjoy.

Finally, we recognize and appreciate Dr. Alan Manning, our faculty highlight this semester, for his contribution of "The Syntax of Plot Frames in Storytelling," an instructive essay analyzing the craft of narrative storytelling. And, of course, we again thank our faculty advisor, Dr. Cynthia Hallen, for her feedback on and approval of this issue.

Stetson Robinson
Editor in Chief

Faculty

The Syntax of Plot Frames in Storytelling

In this article, Professor Manning examines how story frames are made and, by extension, how story frames correlate with syntactic patterns. He then explains how authors that understand this syntactic framing are able to extend this technique to other types of writing.

Alan Manning

Narrative storytelling serves several purposes in and of itself: as a tool for self-expression, as a tool for teaching moral lessons, and as art. But beyond this, storytelling serves as a training ground for the writer, as a place for developing fundamental skills he or she can use to frame forms of writing, which for the beginner are less intuitive and more challenging.

Through the practice of story framing (i.e., creating a plot), writers develop skill in looking past their writing as a linear series of sentences, to see beyond a flat list of separate events to a larger unity in what they write: the overall story frame. In this article we will explore how fundamentally the same syntactic patterns that we find in sentences (in essence, X-bar trees) can be used to describe the larger patterns that organize stories. In later forms of writing, this same syntactic-framing skill can be translated into a rare and crucial ability to organize explanations and arguments in support of unified, larger claims. With this skill, writers find themselves able to do more than tinker with text at the sentence-by-sentence level as unskilled writers tend to do. They are instead able to revise and improve their work at the level of overall organization.

Children's first stories are generally told as a bare sequence of events, without much in the way of illustrative detail. Through the extended practice of narrative (i.e., describing an event sequence), writers develop skill in showing how events unfold with concrete narrative detail, simulating for the reader a perception of the events, real or fictional. In later forms of writing, this developed detailing skill can be translated into an equally rare but valuable habit of using concrete examples to illustrate each and all general explanatory and persuasive statements, re-creating for the reader the writer's own evaluation of the examples.

We will first consider the principles of story framing and then turn to the narrative detail as a way to diagnose student writing development: in other words, the acquisition of storytelling as a specific kind of linguistic structure. We will consider in particular one aspect of narrative detail, the development of character and dialogue, which in other

writing genres translates into a writer's ability to utilize other texts as sources (in research papers, for example) and to consider "audience."

In effect, each successful writer learns to imagine herself or himself, other text-authors, and potential readers as characters debating the meaning of what the writer is writing. Therefore, effective character construction and dialogue is an important developmental step in the acquisition of a mature writing style.

Principles

We begin then with basic storytelling because, unlike advanced forms of writing, most people seem to have an innate sense of the basic form of stories even from childhood (notice too how the child's story emerges in dialogue in the example below):

> Three-year-old: I had a scary dream.
>
> Adult: What happened?
>
> Three-year-old: There were witches, and monsters . . . and crocodiles. And they were scary.
>
> Adult: Were they mean to you?
>
> Three-year-old: No. They were nice. Scary, but nice.
>
> (From the PBS video "Out of the Mouths of Babes")

With this final statement, "Scary, but nice," the child brings this account of her scary dream to a close, and her statement, in a nutshell, shows us the basic upper frame for storytelling. This frame remains quite the same whether it straddles a simple, two-line account like hers, or stretches out to hold together a multivolume epic novel like Tolkien's *The Lord of the Rings*.

What frame could be so flexible? In all genuine stories, complex or simple, a conflict is established by qualitative details, leading the listener/reader to the conflict's resolution. Thus, we can compare the simple dream story to *The Lord of the Rings* using parallel diagrams (see figure 1).

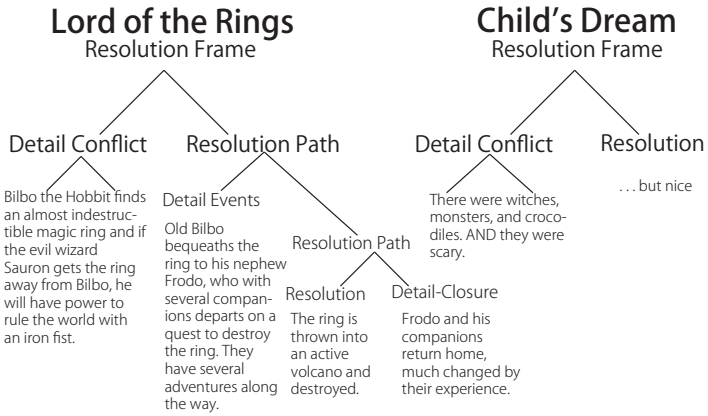


Figure 1. The account of the scary dream and the epic novel *The Lord of the Rings* differ only in length, complexity, and placement of detail leading to resolution of conflict.

As far as this basic framing of conflict and resolution goes, the account of the scary dream and the epic novel differ only in the length, complexity, and placement of detail leading to resolution of conflict. In the most primitive of stories, only enough detail is given to establish the problem; here, the "witches, monsters, and crocodiles" establish the scariness of the dream (the core of the conflict). This problem directs the listener to the resolution (the scary things turned out nice). In the multivolume epic, more details have to be established before the core of the conflict can be understood; Tolkien writes a short novel (*The Hobbit*) detailing how Bilbo the Hobbit finds the magic ring. This short novel in turn establishes the main conflict for the rest of the epic—the problem of keeping the ring away from an evil sorcerer who wants it. This conflict is what motivates the reader to wade through over 1,200 pages, following Tolkien's characters on a quest for resolution (to destroy the ring).

In sum, this two-part framing defines the common agenda for all genuine stories. A conflict frame is established, the conflict having two essential parts (hence the detail-conflict branching fork in the

figure 1 diagrams): providing background details (1) and establishing a core conflict (2). This problematic conflict provokes a listener or reader to seek the resolution frame, which in turn consists of two parts: the detailed conflict already discussed (1+2) and a resolution to the problem (3), as diagrammed in figure 2 (below). More mature stories simply add further detail (1) to the resolution phase (3) of the story. These details create an elongated path of cause and effect. Such resolution paths in mature stories allow the author to lead the reader/listener through several causal event details before reaching the final effect, the resolution.

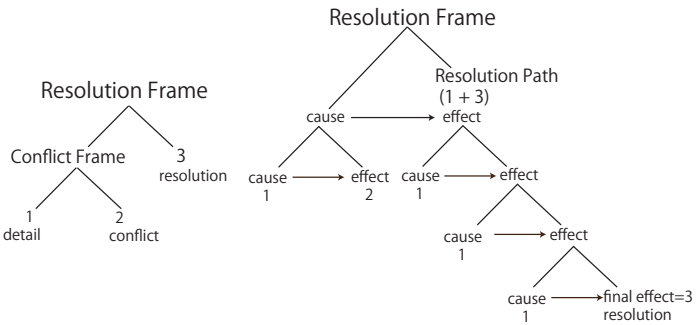


Figure 2. A basic story contains a detailed conflict and a resolution to the problem. More mature stories add details in the resolution phase, which create an elongated path of cause and effect.

Thus, Frodo and his companions have several adventures leading up to their finally disposing of the troublesome ring at the end of the fourth volume. The trick for writers is to recognize one essential upper framework for all genuine stories, despite the different shades and degrees of detail that are found between a child's simplest dream account and an epic novel.

The trick for writing instructors is to take students, with their inborn ability to give a simplistic account of an event, and guide them to frame sophisticated, multilayered narratives. It's not that we even want to turn all students into epic novelists; there's hardly room in the

literary market for the ones we already have. Even so, we certainly do want to bring more students a fair distance in the direction of epic narrative, because the skills developed in narrative framing will prove valuable in framing effective writing in business, technology, and academia; we have far too few effective writers in those markets.

To help students to advance their narrative and story-framing skills, we must be able to determine where their current skills are on the path between the simplest and most complex stories. We've discussed the key features of the first two common stages of narrative framing, from the most basic frame to the more complex path-frame (detailed in figure 2 above). We will now consider the third and most sophisticated of these storytelling formats, the detail-closure frame, where the end of the story is carefully foreshadowed in prior story details, creating the powerful textual effect called closure.

Different nursery stories nicely illustrate the difference between framing a story as a bare cause-effect path and framing a story with illustrative details and genuine closure. The usual bare-path telling of *Cinderella*, *Sleeping Beauty*, or *Snow White*, for example, always includes a merely formal "closure" line, immediately after the actual resolution line, thus:

> **C. Resolution**: . . . because the slipper fit, the prince knew Cinderella was the woman he'd searched for.
>
> **S. W. Resolution**: . . . so it turned out Snow White was still alive. The prince immediately proposed to her.
>
> **S. B. Resolution**: . . . so the prince kissed the sleeping beauty and she woke up.

What follows is in fact the same line each time, the ritual ending:

> **Ritual Closure**: "They got married and lived happily ever after."

Besides this tag line (and the prince), these stories also share a tendency to be told as a sequence of events, each key event being described with just one sentence. Outside of the Disney animated

films, it is rare to find a telling of these stories that has much in the way of descriptive detail. All this is left to the imagination of hearers: what events actually looked like, sounded like, and felt like to the characters.

> **Background Detail**: Cinderella was an orphan who lived with her evil stepmother and ugly stepsisters . . .
>
> **Core Conflict**: They wouldn't let her go to the ball to meet the prince . . .

And so on. There are relatively few nursery stories which are customarily performed, with dialogue and sensory details included. The common exceptions are *The Three Bears*, *Red Riding Hood*, and *The Three Little Pigs*, which typically include short stretches of ritual dialogue:

> "Grandma, what big eyes you have!"
>
> "Little Pig, Little Pig, let me in!"
>
> "All the better to see you with, my dear."
>
> "Not by the hair of our chinny-chin chins!"

Of these, *The Three Bears* is the strongest example. In American culture at least, telling this story also requires ritual dialogue performance in detailing its conflict, its key events, and its resolution. Parents may tell *Cinderella* but they must perform *The Three Bears*. Because more details of the story are made explicit, closure in the story genuinely results from parallels in detail. Because all of the *Three Bears* event-details have parallel structure, the closure effect here is strong and genuine rather than weak and merely formal.

Goldilocks enters the house [**core conflict**]

Event	Bears sit down to eat. (conflict background)	Goldilocks tries the food.	Goldilocks tries the chairs.
Detail 1	"This porridge is too hot."	"This porridge is too hot."	"This chair is too soft."
Detail 2	"This porridge is too hot."	"This porridge is too cold."	"This chair is too hard."

Detail 3	"My porridge is too hot too."	"This porridge is just right."	"This chair is just right."
Outcome	Bears go out for a walk.	She eats Baby Bear's food.	She breaks Baby Bear's chair.

This parallelism of detail continues through the bears' return, their inspection of the house, and their discovery of Goldilocks in Baby Bear's bed. This start-to-finish parallelism eliminates the need for a ritual "happily every after" closure typical of bare-frame fairy tales. Instead, the original conflict (Goldilocks in the Bears' house) is resolved when Goldilocks runs out of the house in fright, and this is the last fixed line of the story.

My mother used to add a final closure line: "And Baby Bear was sad because he would've liked to play with the girl." I have never felt quite satisfied with this ending. Since then, I've thought of a more effective way to follow up on the resolution of *The Three Bears* with further parallel detail. This ending does give final closure, but it requires a small change in the resolution:

> . . . And Papa Bear said, "Somebody's been sleeping in my bed."
>
> And Mama Bear said, "Somebody's been sleeping in my bed."
>
> And Baby Bear said, "Somebody's been sleeping in my bed, and here she is, STILL sleeping in my bed!"
>
> **Resolution**: So the three Bears (being bears, after all, and having missed breakfast) ATE Goldilocks . . .
>
> **Closure**: . . . because SHE wasn't too hot, and SHE wasn't too cold. She was, in fact, JUST RIGHT.

In modern, mature storytelling, the closure effects of parallel detail are not usually this obvious, or this relentless, but they are present, and serve an important unifying function. In *Gone with the Wind* (book and film), for example, Scarlett O'Hara finds herself at the end of the story much changed by events of the Civil War and three bad marriages. However, ironically, Scarlett is stuck in the same emotional situation that she's been in from the first page of the book:

Conflict: Scarlett is in love with Ashley Wilkes, but can't have him.

Closure: Scarlett is (FINALLY!) in love with Rhett Butler, but can't have him.

Now, to bring all of the foregoing storytelling shop-talk to closure, the key differences between a bare-path frame and a detail-closure frame are diagrammed in figure 3.

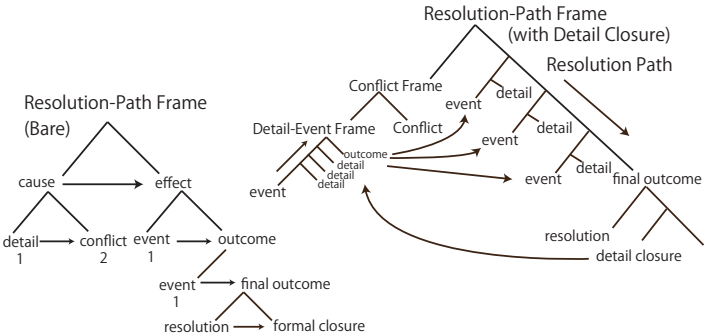


Figure 3. In a fully developed story, the resolution-path frame contains a number of detail-event frames, each consisting of a key event and the detail describing the event.

We recognize mature, detail-closure storytelling as that stage in which the bare, cause-effect sequence (like *Cinderella* the fairy tale) is "fleshed out" with detail (*Cinderella* the motion picture), and that detail is most effective when it unifies the story as a whole with certain parallel elements which create a sense of closure.

As we have already seen, in even the most basic narrative, a storyteller must give at least a little detail (1) to define conflict (2). The dream was scary because of the details: the witches, monsters, and crocodiles. The conflict is necessary so that something may be resolved (3) (the witches, etc., were scary but nice). Thus, the resolution frame contains conflict and resolution; the conflict frame contains background detail and conflict. As shown in figure 3, in a fully developed

story there are naturally a number of detail-event frames, each consisting of a key event and the detail describing the event.

Detail Diagnostics

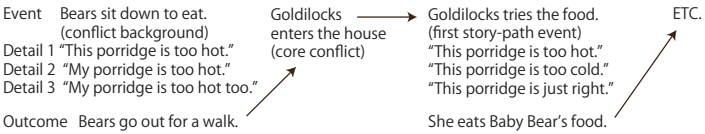
Beginning storytellers may provide just one detail-event frame, enough to establish conflict. Fully-developed storytellers have to learn to balance and manage several detail-event frames to flesh out the bare frame of a story. Skill with detail management also develops through sub-stages, roughly parallel to the overall progression from basic framing (1) to bare cause-effect path framing (2) to detail-closure framing (3). We will now consider these three substages of detail management: the "pure detail" stage (a), the "pure event" stage (b), and the "pure dialoging" stage (c). These stages serve as diagnostics, tools for actually evaluating writers' progress from basic framing to fully mature story framing, full of illustrative detail, which should finally be a balanced mix of each of these pure types of detail.

Recall that any number of events can be stacked to create a resolution path (see figures 1 and 2). That is to say, good stories (well-formed resolution frames) can grow to considerable length (as *Gone with the Wind* and *The Lord of the Rings* do). It is likewise true that any number of details can be stacked to describe the internal shape of an event, from initial event to the event outcome. It will prove important to remember, however, that this event outcome, not the detail stack itself, is what advances the reader or listener another step along the resolution path and what makes the reader or listener feel the story is going somewhere.

Let's go back to *The Three Bears* (just once more, I promise) to compare the different effects of these two methods of bulking up a story: event-stacking vs. detail-stacking.

Novice writers will often first try to lengthen a story by adding pure detail, just as I have lengthened the beginning of *The Three Bears*

below by adding more detail to the first background scene. Notice that I haven't given the Bears many extra things to *do*:



Pure Detail

> Three mangy black bears lived in a little, slightly run-down two-story cottage in the middle of a large, dark wood. It was the first morning in many months that was not bitter cold. It had been a long winter. Pale but welcome sunlight fell across and warmed the rough-hewn wooden table in the little kitchen. Papa Bear's chair, like all the other kitchen chairs, was made of wood and wicker. It creaked and squeaked as Papa Bear sat down. He grumbled a terse and insincere "good morning" at Mama and Baby who were already seated. Papa scratched his left ear absently as he sniffed and stirred the slightly stale-brown porridge Mama had served him.

Now, as "pure" detail, the above passage is all that it should be. I've described the opening scene with sights, sounds, feelings, and even a hint about the smell of the porridge. And yet, if I continue in this style for more than a few lines, most readers will lose interest, and not because they know the story. Although I'm painting a vivid scene here, nothing is happening to hold anybody's interest. On the other hand, a pure sequence of happenings, of cause-effect events that is, will not hold interest either, especially if the sequence comes prior to (or instead of) the main conflict-resolution path that genuine stories require. Then the narrative becomes mere history, a page, as it were, from the "Bear Bible":

Pure Event History

> Abraham Bear begat Isaac Bear; and Isaac Bear begat Jacob Bear; and Jacob Bear begat Judas Bear (and his brethren), and Judas begat Phares and Zara of Thamar; and Phares begat Esrom; and Esrom begat Aram. And Aram begat Aminadab; and Aminadab begat Naasson; and Naasson begat Salmon, and Salmon begat Booz of Rachab; and Booz begat Obed of Ruth; and Obed begat Jesse, and Jesse begat David the king bear; and David the king begat Solomon Bear of her that had been the wife of Urias Bear. (see Matthew 1:2–6)

Readers lose interest in this sort of narrative very quickly indeed. Even if fleshed out by detail, this kind of pure-event narrative can only serve as prologue to the real story (which is exactly what it does in the actual Bible). Only in the last line is there even a hint of something that might be called an interesting detail, because it hints at a social conflict.

Again, this is something that even experienced writers can find themselves doing—writing an overlong historical prologue before they get to anything actually interesting. Attractive writing needs to create some kind of balance between "hot" detail and "cold" event-history. Readers need something in the middle, something "just right" (sorry about that; I'm almost done).

The next version holds interest better, mainly because I've mixed new details with new events, but also because the events I do add are creating a cause-effect path to the story's resolution, to create a full sense of detail closure:

Details + Cause-Effect Events = Resolution-Path Closure

> Mama Bear placed Baby in his highchair, and he began to whimper and fuss.
>
> "Pine nut porridge again?" he bawled mournfully.
>
> "Hush now, Baby," Mama said softly, as she finished ladling the

> steaming goo into Baby's midsize bowl. She put the smallest bowl in front of her own place at the table, dabbing two scant spoonfuls in it. "Papa does the best he can," she whispered, mostly to herself. "Most of the berry bushes have died from the cold, and rabbits and deer are getting hard to find, let alone catch."

In this third version, new events support the new detail, and separate event-details are linked by cause and effect: Baby is upset by the same dull porridge, which he's getting because Papa couldn't find anything but pine nuts to feed his family and this was because the winter was harsh—so harsh, in fact, that the Bear family is even running short on pine nuts, which is why Mama Bear takes so little porridge herself, and this is why her bowl of porridge gets cold before Baby Bear's porridge (in the midsize bowl), which is why Goldilocks eats Baby Bear's porridge all up, which, by the way, is why Baby Bear really ought to eat Goldilocks at the end of the story. (And now I really am done with these poor bears.)

And so we've seen how both details and events are useful and necessary to flesh out a good story, but that either can become an obstacle to effective storytelling unless they are balanced and unless they clearly support the essential frame of a story, from conflict to resolution.

Note that part of the strength of the third passage above lies in my use of dialogue, words that the characters say to one another and to themselves. As narrative detail, dialogue has the special property of being simultaneously an event and a detail (and something else, as we shall see). It's an event because it's something a character does, an act of speech that causes other characters to respond. It's a detail because it's something readers would actually hear if they observed the story events firsthand.

But, like pure sensory detail and pure event history, pure dialogue can also have a third, skewing effect on narrative. This skewing effect is most apparent when storytellers have some specific point they

want to make, some moral, political, or philosophical statement. If they can't think of a way to have characters act out the point, they may choose to have characters discuss the issue:

Pure Dialogue

"I'm against capital punishment," Cosmo said. "Killing a second person won't bring the first one back."

Gerald shrugged. "No, it won't. But what you're also saying is that murderers should be rewarded for being so efficient."

"What're you taking about?" Cosmo asked.

"Well, suppose that crazy old boyfriend of Ellie Mae's came in here in a jealous rage and stabbed you with a knife. You're bleeding bad. But suppose before you die from it, you stumble to the fireplace and get hold of the poker. Whatcha gonna do with it?"

"Bust the guy in the head, of course. Aw, I see where you're going. Sure, I'll kill him first if I can, but that's different, that's self-defense," Cosmo said.

"But you're gonna die anyway," Gerald protested. "Killing Gus won't bring you back, but you'd want to kill him anyway. Now suppose Gus is more efficient with the knife, and gets you right in the heart. You die, and he just gets twenty years, and he's paroled in seven. He's been rewarded for being such an efficient killer. If he hadn't been so efficient, you'd have killed him yourself. What do you think of capital punishment now?"

Cosmo grimaced. "I'm not so sure now."

Now, we've seen how pure detail or pure event narration will inevitably clog or delay the essential story frame of conflict and resolution. Pure dialogue, on the contrary, inevitably creates a bare story frame. That is to say, if people have a genuine conversation, it must have started with some question unanswered for at least one speaker, or some disagreement between two. In a word, there must have been some conflict. Furthermore, the natural outcome of (friendly) conversation is some consensus, some resolution.

Pure detail without action can only create (a) a vivid description (figure 4, left). Pure cause-effect event sequences can only create (b) a history or prologue leading up to the "real" story (figure 4, center). Pure dialogue creates (c) a story frame. Although it will be a fairly bare sequence of utterances, dialogue naturally begins in conflict (a question or disagreement) and ends in resolution (an answer or agreement).

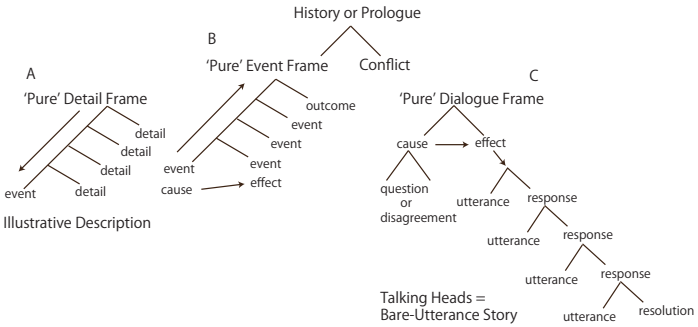


Figure 4. Pure detail without action can only create a vivid description. Pure cause-effect event sequences can only create a history or prologue leading up to the "real" story. Pure dialogue creates a story frame.

Most modern fiction contains stretches of "talking-heads" dialogue like the example above, but some novelists get accused of relying on it too much. They get accused of lecturing their readers, of spoiling their stories with overly pedantic, didactic prose. The reason for this objection has nothing to do with the structure of storytelling, however. Instead, the problem with lecturing dialogue like the last example is that it threatens the basic agenda of literary narrative.

In brief, literature is expected to serve as a substitute for the perception of actual events, and so literary narrative is not supposed to carry with it a fixed, single-minded message. Two people who observe the same event are never compelled by the event itself to draw the same conclusions. What happens when two (or more) people get together and talk about the event is another matter, however. In that

social context, people are inclined to talk about what happened until they come to an agreement.

To the extent that "talking-heads" dialogue persistently drags the reader along to a particular moral, political, or philosophical conclusion, it frustrates reader expectations that a literary text should simulate life itself, full of multiple interpretations. In summary, dialogue detail goes far to satisfy the structural criteria of storytelling, but pure dialogue also overtly leads readers to a particular propositional conclusion and so begins to defeat the usual purpose of modern storytelling, which is entertainment rather than commitment to specific claims.

It's worth remembering, though, that Western philosophy was born in dialogue-heavy storytelling, that is, in the Dialogues of Plato. As philosophy was born, however, the analytic division of discourse genres began. Plato proposed to ban all poets (tellers of fictional stories) from his utopia. Plato's student Aristotle subsequently distinguished poetics (literature), rhetoric (persuasion), techné (craftsmanship), and skientia (knowledge). There are legitimate reasons for this division of genres, but when pushed too far, this division becomes fragmentation: Much of commercial and political rhetoric has become vitriolic; much of technical explanation, unreadable; and much of academic writing, dry and plodding. All in all, the sense of a dialogic negotiation with readers (as characters in a compelling narrative) has been lost.

In future study, we should return to this point and reexamine the common roots of all writing in story and dialogue. Again, the goal here is not to make all writing into an overt dialogue between story characters. Rather, writers may improve their ability in each modern genre by reviewing the basics of storytelling and by consciously evaluating their performance in each genre relative to their storytelling and dialoguing skills, this common center for all writing.

Classroom Practice

Now that we've gone over the basic principles and diagnostic features of storytelling, we are ready to look at some possible techniques for teaching this genre as a foundation for later kinds of writing. Although these storytelling principles I've described are fairly constant, the strategies that teachers might use to implement these principles in the classroom are open and infinite. Readers should consider exercises and assignments that I'll describe in this next section as exemplary suggestions only. Generally speaking, though, teachers and students will want to work through three sequential goals:

1. To be able to IDENTIFY the basic features of story in the writing of others.
2. To be able to OUTLINE the frame for an original story of their own.
3. To be able to DEVELOP a bare story frame with detail, appropriately balanced with cause-effect event sequences and dialogue leading to a satisfying resolution.

Again, generally speaking, teachers and students will want to work through this three-goal sequence with relatively short stories, perhaps only one paragraph or one page in length initially, especially if students have difficulty with longer texts. Once students have mastered the one-paragraph or one-page story, then the whole IDENTIFY-OUTLINE-DEVELOP sequence can be repeated with progressively longer texts, progressively more intricate stories: two to five pages, five to ten pages, ten to twenty pages, and so on.

Practice Identifying Story Features

Humorous anecdotes like those published in *Reader's Digest* serve as abundant, handy examples of the one-paragraph story:

> I began my last semester of college six months pregnant. Walking down the hall in the psychology building one day [BACKGROUND-DETAIL], I was grabbed from behind and spun

> around to face a totally strange young man [CONFLICT]. His eyes fastened on my protruding midriff; his face paled; his books dropped to the floor [DETAIL-EVENT#1]. I shook his arm and asked if he was all right [DETAIL-EVENT #2]. He looked at my face, then whooped and swung me around in a bear hug [DETAIL-EVENT #3]. "Thank God!" he said. "You're someone else!" [RESOLUTION & CLOSURE] (From *Laughter, the Best Medicine*, NY: Berkeley Books, 1982, p. 28)

The basic mechanism of this humorous story (and most others) is that the resolution of its conflict leads readers to a special kind of closure. The punchline leads readers to reinterpret explicitly mentioned details (here, the young man's shock and then sudden joy) and fit these together with unmentioned but now clearly imagined details to create a slightly different but equally coherent story. The original story conflict is explained: the young man grabbed the narrator because from behind she looked like his girlfriend. But this resolution serves as background for conflict in the hidden story: the young man was shocked because he thought his girlfriend was pregnant. This new conflict also has a resolution path: the young man finally looked at the woman's face (EVENT #1). He whooped for joy (EVENT #2) because he was relieved to find the narrator wasn't his girlfriend (RESOLUTION).

This kind of "anecdote exercise" is handy for students because it works at two levels: beginning students might only be asked for the key elements of the first, most obvious story frame; more advanced students might be asked for the key elements of the hidden story created by the punch line.

This sort of anecdote exercise is also useful because it directly illustrates a point I would like to make: that is, story frames (or any other kind of discourse structure) do not physically exist in the words of a written text. Rather, these frames exist in the mind of the writer, whose task is to lead readers to rebuild similar frames in their minds using the text as a guide.

Practice Outlining an Original Story Frame

For most students, this proves to be the most difficult step, to transform their passive comprehension of key story elements into an active production of original work. Once the student has outlined a bare story frame, however, the detail-development process becomes much easier.

The student's first original outline might be a joint effort with the instructor, developed in an office visit or a class-time workshop. Another good first-effort exercise is to have a student pick a common fairy tale and alter it slightly to serve as the frame for a new story. Hollywood screenwriters do this often enough, so there's no reason student writers shouldn't be allowed. But, whether a student begs or borrows his or her first story-frame outline, he or she alone ought to be responsible for the first essential step:

Before they even begin to work out a story frame, writers must discover a key incident—a principle, situation, or event—that serves as the "seed" for the story they will tell. If any story is to be any good, its writer has to care about its essence, so naturally student writers should pick their own key incident to build on. The main problem is that young people often haven't discovered or don't consciously realize that they really care about anything, and so, as writers, this is precisely where they get stuck.

To get around this problem, I tell students in my introductory linguistics class that their main task in their story-writing assignments is to illustrate some concept or principle they've learned about in class, whatever idea they've found most interesting, but preferably some idea they've seen actually exemplified in real life, outside of class.

So, for example, after one of my lectures on the differences between animal communication and human language, one of my students was particularly struck by the idea that even a super intelligent dog would not be able to speak, since the animal would lack a human vocal tract specifically evolved for resonating human vowels and articulating human consonants.

This idea led this student to her key incident: she imagined the frustration that a superintelligent animal would suffer if it knew about language but was unable to use it. All she had to do then was pick a preexisting story frame such as *Red Riding Hood*:

> **Conflict**: Red Riding Hood tells a wolf she's taking goodies to Grandma (BIG MISTAKE).
>
> **Event 1**: Wolf takes shortcut to Grandma's.
>
> **Event 2**: Wolf eats Grandma and puts on her outfit.
>
> **Event 3**: Wolf tricks Red Riding Hood into coming close to the bed.
>
> **Resolution**: Wolf eats Red Riding Hood (punishment for little girls who talk to strangers).

With a key incident of her own in mind, the student was then able to reshape this familiar story frame into something original, creative, and instructive:

> "Hello, my name is Maria Lopez, but my friends call me Little Red Riding Hood," Maria said to the wolf she met in the woods.
>
> The wolf glared at the little girl. He hated it when humans tried to talk to him. Try as he might, he could not make his carnivorous wolf mouth form the words that he heard so clearly in his head. This caused him great frustration and was consequently a touchy subject.
>
> Maria decided that the reason behind the wolf's silence was that he couldn't understand English. She tried once again to speak to the wolf.
>
> "Hola, me llamo Maria Lopez, pero mis amigos me llaman Capulacita Roja."
>
> The wolf grew ever more furious with the little girl and something snapped inside his head. Meanwhile, the little girl decided that the wolf's silence was purely rude.
>
> "I don't understand why you won't talk to me. I don't think you're very nice. Well, I'll be on my way then over to my grandmother's house."

> The wolf was quite annoyed with Little Red Riding Hood and proceeded to plot out revenge for the little girl's insensitivity . . .

The key incident students invent in a fairy-tale setting may become the background for their story's main conflict (as in the example above). This is perfectly acceptable, but more mature writers typically work backward from a key incident they have directly experienced themselves, developing a separate conflict and resolution path that leads up to the key incident, then past it to a resolution of the original conflict. This approach indicates that students are planning ahead and thinking in terms of concrete examples, but they still must watch for plotting pitfalls.

For example, many students have approached me with proposed story outlines based on my linguistics lecture on the difference between literal, direct language ("Hand me the salt shaker right now") and more socially acceptable, indirect language ("Could you pass me the salt, please, when it's convenient?"). This linguistic principle resonates with many precisely because we've all experienced the different emotional effects created by direct and indirect language. On several different occasions, different students have proposed to me this very same story frame:

> **Conflict**: The main character is accused of being socially inept because he uses blunt, direct language when he ought to have been indirect (polite, tactful).
>
> **Event 1**: Character expresses wish that everybody would learn to be completely honest and direct in what they say, just like him.
>
> **Event 2**: Character has dream where he gets his wish. Everybody is blunt and direct, and the character's feelings are hurt.
>
> **Resolution**: Character wakes up and realizes his mistake.

As popular as the "dream-sequence" story frame is with TV sitcom writers, this is one plot device that I try to discourage, and not just because it's been beaten to death. The main reason I try to get students

to improve on this outline is that it automatically suppresses the authentic, concrete experiences that students have had, experiences that made them interested in direct and indirect language to begin with. The next version is much improved:

> **Conflict**: Boyfriend breaks up with main character. She is hurt because he has used several "break-up" clichés meant to spare her feelings: "I need some space," "You deserve better than me," etc.
>
> **Event 1**: Main character complains to roommate that she wishes boyfriend would just tell the unvarnished truth: he's met somebody else he'd like to date more than her.
>
> **Event 2**: Wise roommate (a linguistics student of course!) later makes several completely direct, completely true remarks to main character: "Those jeans make your behind look huge," and "I need your dictionary for a paper I'm writing. Give it to me," and so on.
>
> **Event 3**: Main character is outraged.
>
> **Event 4**: Roommate explains that she was just trying to show how important indirect language is to happy human relations.
>
> **Resolution**: Main character forgives ex-boyfriend and gets on with her life.

In helping students develop fiction-writing skills, I encourage them to work through and then move beyond fairy-tale and dream-sequence story frames, to finally plot a story as close to reality as possible, but still fictionalized as much as necessary. Writers make good use of fictional story frames to protect the identity of actual people (including themselves!), avoiding, among other things, embarrassment and libel suits if anybody actually reads their stories. To guarantee that people stay interested in what they write, writers also must edit and reshape real events until they fit naturally with the conflict-resolution frame of storytelling.

I make a final point of this because, once they're converted to the idea of a realistic story, some writers want to insist on a direct, completely accurate account of their personal experience, with all the

messy, even annoying details intact. At this suggestion I point out that telling an experience "exactly as it happened" entails many of the same dangers we encounter in being totally honest and direct in what we say to other people. We run a great risk of boring readers into a stupor with all our actual thoughts and experiences, or we run the still greater risk of offending them with the unvarnished truth.

Practice Developing Interesting, Goal-Driven Detail

Once writers have worked out an adequate story frame, that has grown out of some key incident they are genuinely interested in writing about, they can begin the main business of fleshing out a story with detail. As noted in the diagnostic section of this chapter, we can expect some beginners in their first efforts to get bogged down in descriptive detail.

> I woke up and heard the birds chirping. It was nice to hear the sound of birds chirping over the sound of the traffic. My first reaction was that it was going to be a nice day outside, but then I remembered that those same Russian birds were there even when it was thirty degrees below zero. My roommate had shut all the windows before she went to sleep. This didn't exactly leave much oxygen in the room. I woke up suffocating and faced a new set of cockroaches.

Still others we can expect to pass over too many events, sometimes without detail, but more typically without getting to the main conflict of the story quickly enough:

> When I was first appointed general over the Nephite army, there had already been many wars and contentions over many years. Many of the battles and wars were fought because of the spirit of contention that lived in the hearts of our enemy's forefathers. They felt they had been robbed of their right to rule over us and they then passed the spirit of contention on down

> to their sons and daughters who then passed it on to their sons and daughters. But there is also another reason why the great battles of my career came to pass . . .

Even experienced writers tend to write longish detail or historical prologues in their first detailed draft. Naturally, then, the writing student's most important remedial exercise at this detail-development stage is to search through their draft until they come to the main conflict in their story, the problem that gets solved by the end. I then instruct writers to delete everything prior to that main conflict and start the story at that point.

If the student's narrative contains no conflict that ever gets solved, he or she needs to go back to the story-outlining stage described above. Usually though, there's something to work with. Let's rejoin the first example above, several dozen lines later in the actual narrative:

> I was sitting on the bus wrapped in so many clothes. It was thirty below that day and my babushka [landlady] had made sure that the last coat I put on over the rest of them was a very Russian-looking purple coat. I brought my English/Russian flash cards along with me. [MORE DETAIL DELETED] I hadn't really paid any attention to the man sitting next to me until I became aware that he was staring at my flash cards. [MORE DETAIL DELETED]. He pointed to my flash cards and asked if those were English words next to the Russian, and I responded, "Da." I was then prepared for him to start asking the usual, insinuating questions: Was I American? Would I ever date or marry a Russian man? Instead he asked me if I understood English very well and if it was a hard language to learn. He thought I was Russian, all wrapped up in that ugly purple coat as I was. If I could only keep him thinking I was Russian until I could get off the bus, I thought, maybe I could avoid (for once) being hit on by this guy. Trouble was, my accent would give me away unless I kept saying "Da" to everything he said.

The passage that remains after all this cutting still establishes the most essential background details: the narrator is an American girl in Russia, on a bus, wrapped up in heavy coats against the cold. Only these details are needed to establish the main conflict, a difficult choice this girl faces: to give up her protective "Russian" appearance and be harassed, or to keep saying "yes" to anything this stranger might say . . . or ask.

Students and mature writers alike have to find the courage to delete all the clutter, scraps of imagery, and unconnected thoughts obscuring an otherwise interesting and effective story. Details and events must either support the main conflict or lead to the final resolution. To do otherwise is to do something other than tell an effective story.

Professor Manning received his BA in linguistics from Brigham Young University and his PhD in linguistics and technical writing from Louisiana State University. He began teaching at BYU in 1994 and is currently the advisor for the student journal Leading Edge *and the area coordinator for the editing minor. He has also written two novels,* Supposition Error *(1996) and* Love Story Logic *(2000).*

Symposium

The Acceptability of Linguistic Back-Formations

Traditionally, it has been believed that educated users of English largely disregard the existence and function of back-formations (lexemes that are created from an already existing word) as words. The author examines this notion by studying the acceptability of back-formations among English speakers.

Brooke Barton

Back-formations have been creeping into the English language for decades and some have even become widely accepted parts of the lexicon. Scholars "who have dealt with English word-formation . . . have all included some discussion of coining words by . . . back-formation" (Kreidler 1979, 24). A back-formation is "the formation of a new lexeme by the deletion of a suffix, or supposed suffix, from an apparently complex form by analogy with other instances where the suffixed and non-suffixed forms are both lexemes" (Bauer 1983, 64). Generally, back-formations are not considered to be legitimate words by scholars, by academics, or in dictionaries, but are back-formations in English becoming more accepted?

There are numerous examples of back-formations in the English language, and many are considered incorrect from an academic perspective. Others have become accepted over time through repeated and common usage. The verb *laze* is an example of a back-formation that remains unaccepted by scholars. It does, however, appear in the *Oxford English Dictionary*, albeit with very seldom usage over the past three centuries. *Laze* comes from the word *lazy*, and is analogous with the pair of words *craze* and *crazy*. Examples of back-formations that have crept into the English language and have now become accepted and completely legitimate are the verbs *edit* and *exhibit*, which historically did not exist in the English language. Both are back-formations that came from their original noun counterparts *editor* and *exhibitor*, which describe people (Bauer 1983, 230). The back-formations came about when the roots of these words were assumed to be verbs, describing what *editors* and *exhibitors* do, which would be to *edit* or to *exhibit* something. Back-formations "may come into existence by pure invention, but this is rare. . . . More typically, new words are fashioned from elements already present in the language" (Kreidler 1979, 25).

The "great majority of back-formations in English are verbs" (Bauer 1983, 230). Verbs constitute 87 percent of all back-formations (Pennanen 1966, 119). When a back-formation forms, it is generally because an affix attached to some root has brought about the

psychological conception of the perceived verb. This affix is assumed to have been attached to the verb, making it into a noun, adjective, or any other part of speech. In many historical cases, this assumption was incorrect because the noun existed without an original verb counterpart. New verbs were created in the English language as the result of back-formations, as evidenced by the aforementioned examples of *edit* and *exhibit*. These words began to be used, then became widely accepted, and eventually became legitimate according to scholars, academics, and dictionaries.

The number of back-formations in the English language is ever increasing. There are verbs that come from adjectives: *adolesce* from *adolescent*; *luminesce* from *luminescent*; *drowse* from *drowsy*; *enthuse* from *enthusiastic*. Verbs can also come from nouns: *caretake* from *caretaker*; *orientate* from *orientation*; *demograph* from *demographic*; *quantitate* from *quantity*. In their study on back-formations in both the Dutch and English languages, Booij and van Marle found usage in America of the words *anthologize*, derived from *anthology*; *ventriloquize*, derived from *ventriloquist*; and *oppone*, derived from *opponent* (1989). None of the preceding verb examples are academically accepted as "real" words, although they do appear with a few usages in the *Oxford English Dictionary* and the *Corpus of Contemporary American English*. Since back-formation verbs like these have documented usage, they evidently have psychological reality. I hypothesized that back-formations are becoming more accepted in English, even if they are not considered standard by academic sources. I constructed a questionnaire to test this research question.

Back-Formations Are Accepted

Back-formations are accepted as valid words in English by everyday speakers—even if prescriptivists, scholars, or academics do not recognize them as legitimate. Back-formation "is to be expected" and "emerges naturally" (Plag 2003, 186) in everyday, productive, spoken

language. This idea is the alternative hypothesis and the one this study aimed to validate. In examining the ways in which new words come into a language, Kreidler stated that "new words come into existence . . . rarely, and they come in a limited number of ways" (1979, 25). The process of back-formation is one of the highest contributors of new words in a language, and indeed one of the "limited" ways new words come about, as defined by Kreidler. He concluded that back-formation not only exists in the English language but that words resulting from back-formation are psychologically processed and understood as legitimate words, regardless of whether they were found in a dictionary or explicitly learned. As long as a word has some recognizable and existing root, or was "formed productively by means of a process of analogical back-formation" (Spencer 1988, 679), an English speaker will generally accept back-formations without questioning their validity. In his defense of back-formations, Kreidler stated that "in our daily use of language we produce and we process numerous utterances. . . . Production is, in general, a constant creation" (1979, 25).

Part of this production is drawing from our knowledge of **existing** words and creating new parts of speech from them to fit a **meaning**, context, or situation in which they are needed. For example, in describing some contraption that broke and needed to be repaired, a productive speaker or writer might draw an analogy and produce the back-formation verb *contrapt* to fit the meaning he or she needed to express. A sentence spoken by an English speaker in everyday language, such as "The contraption he made yesterday broke, so he needed to contrapt something new," would not confuse the majority of English speakers. Back-formation verbs, such as *contrapt* in this instance, are not frowned upon or explicitly corrected but rather are used and accepted in English.

If the alternative hypothesis is supported by this study, and the null hypothesis rejected, a possible implication would be that our grammar is rule governed, but that our lexicon is not finite. This would mean we can attach or detach morphological affixes from existing words to

create new, unreal words that still carry understood and psychologically valid meaning. In everyday, productive written or spoken English, our lexicon is not limited to a finite set of memorized, learned words. If the brain knows a word exists, it can create analogical links to other words that result in the production of back-formations. These words will be accepted and understood by an English speaker.

Back-Formations Are Not Accepted

While back-formations might carry meaning if understood and applied in context, they are not preferred in English nor are they readily used and accepted. If this null hypothesis fails to be rejected, it implies that there is not a significant proportion of English speakers who would accept back-formations as legitimate words without questioning them or seeking further clarification of meaning. Furthermore, people who would not accept a back-formation would rarely, if ever, produce back-formations themselves. As an example of how an English speaker might view back-formations in an unfavorable light, take into consideration a blog post from Stan Carey about his distaste for back-formations and his opinion of their unacceptability. In discussing those who speak relaxed English, he wrote that back-formations are "irregular, even ignorant, and suitable only for informal use in slang or jokes." He further stated that "there is no need for them because the semantic niche they purport to inhabit has already been filled" (Carey 2009).

The null hypothesis discounts back-formations as readily accepted in the English language. Back-formations can be attributed to "untrained" eyes and ears that believe "any combination of a root and one or more affixes" can be broken up into smaller parts. English speakers do not use back-formations as real meaningful words, but "some can be used for comic effect." This is done intentionally and not as part of productive language. In opposition to the hypothesis of productive

language, the null hypothesis asserts that when words "are new, they'll strike many people as odd . . . [and] still struggling for acceptance" (Lynch 2008, 199). Back-formations are "word[s] invented by subtracting an imagined affix" and would never be used without deliberate thought in English, because they go against an innate sense of correctness and generally have a "low level of acceptance" (Lyman 2007). Back-formations are not used in productive language, but rather conjured up and deliberately invented to fill a specific need, such as comically portraying an uneducated individual.

The implications of the null hypothesis, if it fails to be rejected, include that as English speakers we have a finite, unproductive set of words in our lexicon, and our brains are not productively making analogical connections to create new words when the need arises. We only use words that we have heard before, have seen somewhere, or have been taught are legitimate. The words that presently constitute our lexicon are the words we will use, and we are slow to accept new variations of these words, let alone deem them acceptable.

Methodology

I measured whether or not test subjects found a back-formation acceptable. The independent variable was back-formations and the dependent variable was the acceptability of each back-formation. I was not concerned with whether or not the subjects knew if a word is considered academically correct or if it is in a dictionary. I was attempting to generate descriptive data, not prescriptive results. The participants in this study completed an internet questionnaire that I had created and invited them to participate in. The participants and the questionnaire are described in more detail later on.

The null hypothesis, H0, was that back-formations are not accepted in English. The alternative hypothesis, HA, was that back-formations are accepted in English. My threshold of significance was

defined in terms of a 95 percent confidence interval. I set up an Excel spreadsheet that was programmed to calculate the perceived "weirdness" of each back-formation, or in other words, its acceptability, according to my results. Each word was entered into the spreadsheet and its proportion of judged weirdness determined. For example, the back-formation *enthuse* was seen 64 times by participants. Of the 64 people, 16 judged this verb as "weird." In other words, 16 people would not accept it as a legitimate word. The number 16 was put in cell B6 on the Excel spreadsheet, and the number 64 in cell D6. The proportion of this word's weirdness was therefore 25 percent (16/64), calculated in cell F6 of the Excel spreadsheet, which was programmed with the formula B6/D6. A 95 percent confidence interval was then created in cells G6 and I6 of the Excel spreadsheet. The cells were programmed with the following confidence interval formulas, respectively: =IF(F6-1.96*SQRT(N6*N6/(D6+4)) <0,0,F6-1.96*SQRT(N6*P6/(D6+4))); =IF(F6+1.96*SQRT(N6*P6/ (D6+4))> 1,1,F6+1.96*SQRT(N6*P6/ (D6+4))). For the word *enthuse*, the Excel spreadsheet created a confidence interval of 15 percent to 35 percent, meaning that it can be stated with reasonable (95 percent) confidence that between 15 percent and 35 percent of all people will judge this word as "weird." The threshold of significance is that if the lower end of the confidence interval is 50 percent or more, a majority will find the word weird, and therefore unaccepted. If the upper end of the confidence interval is less than 50 percent, there is evidence of a minority and people do not find the word weird, but accept it. Referring back to *enthuse* as an example, this word can be classified as acceptable because the upper end is less than 50 percent, so only a minority find it "weird."

Using confidence intervals as my threshold of significance, I calculated the acceptability of each individual word but also of all the back-formations as a collective whole. The proportion of perceived weirdness for all back-formations combined was inserted into the Excel spreadsheet and a total confidence interval calculated.

Subjects

There were 357 total participants in this study: 92 males (25.8 percent), 263 females (73.7 percent), and 2 individuals who failed to indicate their gender on the questionnaire (0.5 percent). After creating the survey, I sent an invitation to participate to everyone that I knew via e-mail and the social networking site Facebook.

All participants were aged 18 and older. People ages 17 and under were considered a vulnerable population and were not permitted to proceed any further in the questionnaire. As is seen in the table below, the majority (186/357; 52.1 percent) of participants fell in the age range of 18–24 years. The remaining 47.9 percent of participants (171/357) were over 24 years of age.

Age Group	**Responses**	**Percent**
0–17	7	2%
18–24	186	52%
25–29	29	8%
30–35	24	7%
36–44	25	7%
45–55	35	10%
56+	51	14%
Total	**357**	**100%**

Figure 1. Shows the specific age demographics.

I took into account both native and nonnative speakers of English in this study, although the overwhelming majority (337/357; 94.4 percent) of participants were native speakers. Only 19 participants (5.3 percent) were nonnative speakers, and one individual (0.3 percent) is unaccounted for because he or she failed to indicate his or her language status on the questionnaire.

The participants came from varying educational backgrounds. The majority of participants (324/357; 90.8 percent) had some level of college education, ranging from some college to a PhD. The other

33 participants (9.2 percent) declared an education level of some high school, high school graduate, or other. Participants were allowed to select from eight options on the questionnaire, as shown below.

Education Level	**Responses**	**Percent**
Some high school	3	1%
High school graduate	27	8%
Some college	143	40%
Associate's degree	30	8%
Bachelor's degree	108	30%
Master's degree	36	10%
Doctoral degree	7	2%
Other	3	1%
Total	**357**	**100%**

Figure 2. Shows specific education level demographics.

Stimuli

The acceptability of back-formations was tested using a Qualtrics survey I created. A total of 75 words were used in this Qualtrics survey: 25 back-formations (BF), which I analyzed; 25 real but less common words (RLC); and 25 nonreal words (NRW), both of which were distracter words I did not analyze. These words were randomly selected from a variety of resources: the *Oxford English Dictionary*, the *Corpora of Contemporary and Historical American English*, Wikipedia, database journal articles, and published books. A list of all 75 of these words by category is available in appendix A.

Using these 75 words, 25 sentences were created, each containing three words. These 25 sentences were further divided into five blocks containing five sentences each, with the following combinations of words: 1 block (5 sentences)—1 BF, 2 RLC; 1 block (5 sentences)—2 NRW, 1 RLC; 1 block (5 sentences)—2 NRW, 2 BF; 2 blocks (10 sentences)—1 BF, 1 RLC, 1 NRW. All 25 sentences were

present in the questionnaire, but they were randomized so that each participant only saw five sentences total—one from each block. The words were shown in sentences to give them context. The three words paired together in each sentence were randomly assigned, and I created the sentences with the objective of giving all words an equally authentic context. All 25 sentences are available in appendix B.

Each of the five sentences a participant saw was accompanied by the prompt *Please select the words that sound weird to you*. The word *weird* was intentionally undefined on the questionnaire, leaving it open to interpretation for participants to perceive the word as "weird" in any context, grammatical or lexical. Participants could run their computer mouse over each word in the sentence, and the three words in question would become selectable as the mouse passed over them. Participants were told before beginning the questionnaire that if they clicked on a word, it would be marked as "weird." Upon clicking on a word, the word became highlighted green. None of the participants knew what was being tested. Thus, they did not know that a word they were seeing came from one of three categories, that two-thirds of the words were distracters, or that only one-third of the words (i.e., the back-formations) were actually being studied. Additionally, the participants were not in a controlled environment. They completed the questionnaire at any location of their choice, on any computer, and at any time.

Results and Analysis

To summarize the results, back-formations had medium acceptability when compared with other words (see figure 3). Real, less-common words had the highest acceptability. They were judged as the least "weird"; generally between zero percent and 24.9 percent of people marked them as "weird." Nonreal words had the lowest acceptability. They were considered the most "weird," occupying the upper range of between 86.3 percent and 96.7 percent proportion judged weirdness.

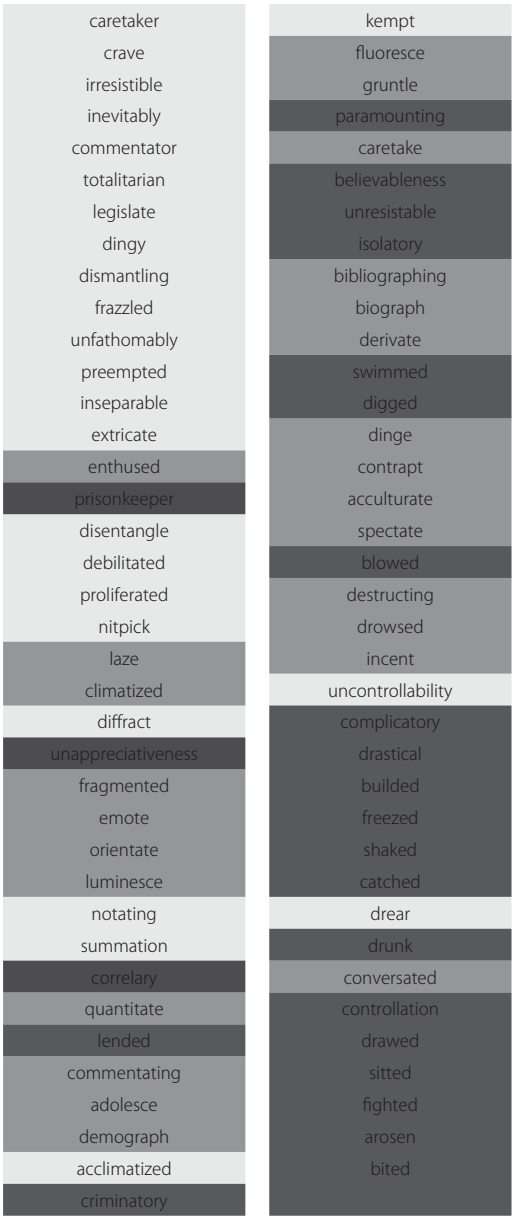


Figure 3. All 75 words, arranged on a horizontal axis by the proportion of people who judged them "weird," from least weird to most weird. Disregard the small-print words and numbers; this figure is intended to show, generally speaking, that real, less-common words (light gray) had highest acceptability, back-formations (dark) had middle acceptability, and nonreal words (darkest) had the least acceptability.

The proportion of weirdness between 32.2 percent and 80.7 percent consisted almost solidly of back-formations.

The least weird back-formation, as judged by participants, was *enthuse*, with 25 percent of people marking it as "weird." The weirdest back-formation, as judged by participants, was *conversate*, with 92 percent of people marking it as "weird." Three words that fell right in the middle out of the total 75 words were all back-formations: *emote* (49.4 percent judged weird), *orientate* (50 percent judged weird), and *luminesce* (50.9 percent judged weird).

As stated previously, the level of significance was determined by a confidence interval calculated for each individual back-formation. If the lower end was 50 percent or larger, there was evidence of a majority judging the word as "weird." If the upper end was less than 50 percent, there was evidence of a minority judging the word as "weird." By this definition, only three back-formations out of 25 (12 percent) were found acceptable. These three back-formations and their confidence intervals were *enthuse*, 15–35 percent; *laze*, 21–44 percent; and *climatize*, 25–46 percent. The other 22 back-formations (88 percent) were judged as weird.

All 25 back-formations were seen by participants a total of 1,654 times. Out of this number, participants marked the back-formation as weird 1,048 times. Thus, the proportion of back-formations judged as weird was 63.4 percent (1048/1654). The confidence interval generated for this cumulative total was 61–66 percent. This interval suggests that the majority of participants found back-formations weird and not acceptable, because the lower end is well above 50 percent.

Below are several tables that break down the results by demographic. These results will not be discussed at great length in terms of statistical significance because only one of them could be considered statistically significant according to the methodology (see figure 8). In terms of proportions of people who marked each back-formation as "weird," it is interesting to note that there are no striking differences between any groups in any category. No trend, such as one gender being drastically more accepting of back-formations than the other, exists for

"Weird"	Out of	Proportion	BF
16	64	0.250000000	enthused
19	59	0.322033898	laze
26	73	0.356164384	climatized
27	55	0.490909091	fragmentated
42	85	0.494117647	emote
33	66	0.500000000	orientate
27	53	0.509433962	luminesce
35	67	0.522388060	quantitate
34	59	0.576271186	commentating
35	60	0.583333333	adolesce
34	57	0.596491228	demograph
54	82	0.658536585	fluoresce
44	66	0.666666667	gruntle
45	67	0.671641791	caretake
47	67	0.701492537	bibliographing
43	60	0.716666667	biograph
53	73	0.726027397	derivate
43	57	0.754385965	dinge
62	82	0.756097561	contrapt
60	77	0.779220779	acculturate
52	66	0.787878788	spectate
53	66	0.803030303	destructing
49	61	0.803278689	drowsed
46	57	0.807017544	incent
69	75	0.920000000	conversated

Figure 4. The 25 back-formations in a table showing the proportion of participants who marked each back-formation as "weird," descending from least weird to most weird. The back-formation *conversate* was an outlier, as most of the back-formations generally fell between a 32.2- and a 80.7-percent proportion of judged weirdness.

any demographic. Generally speaking, being in any one group above another does not guarantee one is more or less likely to find back-formations acceptable. This statement is true across all categories.

Every proportion in the tables below falls between 60.12 percent and 66.17 percent with the exception of two outliers (see figure 8). This supports the trend previously identified: back-formations are not accepted but are not rejected either; they seem to have middle acceptability (see figure 3). The average of all the proportions on the tables below is 62.54 percent (10.01/16). Removing the two outliers, the average is 63.82 percent (8.96/14).

Gender	**"Weird"**	**Out of**	**Proportion**	**Conf. Int.**
Males	294	445	0.660674157	62–70%
Females	765	1229	0.622457282	60–65%

Figure 5. Gender comparison of the perceived "weirdness" of back-formations.

Age group	**"Weird"**	**Out of**	**Proportion**	**Conf. Int.**
18–24	536	863	0.621089224	59–65%
25–29	83	131	0.647058824	55–72%
30–35	72	110	0.654545455	57–74%
36–44	64	104	0.615384615	52–71%
45–55	108	165	0.654545455	58–73%
56+	144	231	0.623376623	56–69%

Figure 6. Age comparison of the perceived "weirdness" of back-formations by age group.

Speaker	**"Weird"**	**Out of**	**Proportion**	**Conf. Int.**
Native	956	1524	0.627296588	60–65%
Nonnative	51	80	0.637500000	53–74%

Figure 7. Comparison between native and nonnative English speakers of perceived "weirdness" of back-formations. It is important to note that because such a small number of nonnative speakers responded to the survey, these results are likely insignificant.

Education	"Weird"	Out of	Proportion	Conf. Int.
High school graduate	67	114	0.587719298	50–68%
Some college	419	643	0.65163297	61–69%
Associate's degree	90	136	0.661764706	58–74%
Bachelor's degree	300	499	0.601202405	56–64%
Master's degree	111	169	0.656804734	59–73%
Doctoral degree	14	29	0.482758621	31–64%

Figure 8. Comparison between the perceived "weirdness" of back-formations by education level. In this table lies the one result that could be considered statistically significant according to the methodology: the confidence interval for *Doctoral degree* is 31–64 percent. This suggests that people of the highest education level do not find back-formations weird. While this idea is intriguing, only 7 people who took the survey fell in this category (7/357; 1.9 percent). More participants of this education level would be needed to prove anything significant about this statistic or to accurately compare it to other statistics. On the Qualtrics survey, there were two additional options available that are not summarized in the table above: *Some high school*, listed before *High school graduate*; and *Other*, listed after *Doctoral degree*. There were only 3 people who fell into each of these categories (3/357; 0.84 percent); therefore they were not significant enough to be calculated.

Conclusion

According to my methodology, the results would be significant in showing that back-formations are accepted if the cumulative total of all back-formations fell in a confidence interval with an upper end of less than 50 percent. However, the upper end of the calculated confidence interval for this study was 66 percent; therefore, the null hypothesis fails to be rejected and the alternative hypothesis is not supported. In other words, according to this study, back-formations are not becoming more accepted in the English language, but rather are still considered weird by the majority of people.

The main implications of this conclusion, as discussed previously, are that English speakers have a finite set of words in their

lexicon and are not productively making analogical connections to create new words when the need arises. English speakers will more readily use words that have previously been heard, seen, or taught as legitimate. The words that presently constitute the typical English speaker's lexicon are the words that will be used, and speakers are slow to accept new variations of these words.

There were several limitations to the study pertaining to variables and factors that were out of reasonable ability to control. Each participant's area of study or expertise was not identified; therefore, some participants likely had an advantage in completing the questionnaire over others (a linguist compared to a chemist, for example). The participants' thoroughness with or attention given to the survey was not accounted for, nor was their environment or the time of day at which the survey was taken controlled. The survey had no time limit, and the time participants took to complete it ranged between 2 and 19 minutes. Another limitation was how many times each back-formation was actually seen. Because the survey was randomized and distracter words were included, the back-formations were seen by participants anywhere from 44 to 86 times.

Future Work

As mentioned above, each back-formation was seen between 44 and 86 times. In a future study, this range could be decreased by having 10 sentences instead of 25. There would therefore be 30 words total instead of 75, with 10 of those words as back-formations and 20 as distracters (10 RLC and 10 NRW).

It would be interesting to do an exclusive comparison between native and nonnative speakers in a future study, doing research beforehand and hypothesizing on nonnative speakers' acceptance of back-formations. For example, do nonnative speakers have a lower acceptance of back-formations because they have been explicitly taught vocabulary, whereas many native English speakers have not?

A future study could compare only two age categories: those aged 17 and under to those 18 and over. This could be an interesting commentary on a potentially strong division between teenage and adult perceptions of language. Do younger people show more acceptability for back-formations? Does the level of education after high school have a significant effect?

Testing back-formations in other contexts would likely produce different results from this study and would provide interesting comparisons. For example, testing back-formations for both their written and spoken acceptability would be interesting. Are back-formations perceived as less "weird" in a spoken context? Does orthographic form have anything to do with perceived "weirdness"? Another context in which to test the back-formations could be no context at all—in isolation, without the aid of a constructed sentence to give meaning. This would provide more insight into the lexicon and vocabulary knowledge of people rather than their ability to form analogical connections and understand productive language in context.

References

Bauer, Laurie. 1983. *English Word-Formation*. Cambridge: Cambridge University Press.

Booij, Geert & Jaap van Marle. 1989. *Yearbook of Morphology*. 2nd ed. Rhode Island: Foris Publications U.S.A.

Carey, Stan. 2009. "Back-Forming Back-Formations." Sentence First. <http://stancarey.wordpress.com/2009/04/28/back-forming-back-formations>

Davies, Mark. 2008. *The Corpus of Contemporary American English* (COCA): 400+ million words, 1990-present. <http://corpus.byu.edu/coca/>

Hall, Robert A. 1956. "How We Noun-Incorporate in English." *American Speech* (31) 2:83–88.

Kreidler, Charles W. 1979. "Creating New Words by Shortening." *Journal of English Linguistics* (13) 1:24–36.

Lyman, Darryl. 2007. "Back-Formations: Words Invented by Subtracting Imagined Affixes." *Associated Content.* <http://www.associatedcontent.com/article/351939/backformations_words_invented_by_subtracting.html>

Lynch, Jack. 2008. *The English Language: A User's Guide.* 2nd ed. Newburyport: Focus Publishing/R. Pullins Company.

Oxford English Dictionary. 1989. *OED Online.* 2nd ed. Oxford: Oxford University Press. <http://www.oed.com/>

Pennanen, Esko V. 1966. *Contributions to the Study of Back-Formation in English.* 9th ed. (4) A. Tampere: Acta Academiae Socialis.

Plag, Ingo. 2003. *Word-Formation in English*. Cambridge, United Kingdom: Cambridge University Press.

Spencer, Andrew. 1988. "Bracketing Paradoxes and the English Lexicon." *Language: Journal of the Linguistic Society of America* (64) 6:663–682.

Appendix A

Back-formations *(BF)*	**Real, less-common words** *(RLC)*	**Nonreal words** *(NRW)*
1. laze	1. crave	1. prisonkeeper
2. fragmentate	2. irresistible	2. builded
3. conversate	3. inseparable	3. swimmed
4. quantitate	4. totalitarian	4. blowed
5. biograph	5. kempt	5. isolatory
6. dinge	6. caretaker	6. complicatory
7. adolesce	7. frazzled	7. shaked
8. emote	8. disentangle	8. correlary
9. enthuse	9. nitpick	9. unresistable
10. gruntle	10. inevitably	10. drastical
11. spectate	11. unfathomable	11. controllation
12. luminesce	12. drear	12. lended
13. derivate	13. uncontrollability	13. drunk
14. drowse	14. diffract	14. drawed
15. acculturate	15. dingy	15. arosen
16. climatize	16. preempt	16. unappreciativeness
17. commentate	17. proliferate	17. believableness
18. demograph	18. acclimatize	18. fighted
19. incent	19. debilitate	19. digged
20. fluoresce	20. extricate	20. sitted
21. contrapt	21. commentator	21. bited
22. destruct	22. legislate	22. paramounting
23. orientate	23. notate	23. freezed
24. bibliograph	24. summation	24. criminatory
25. caretake	25. dismantle	25. catched

Appendix B

Block 1 (5) 1 BF; *2 RLC*

1. Whenever I laze around on the beach, I *crave* piña coladas and find a cool dip in the ocean *irresistible.*
2. Even a seemingly strong and *inseparable* people like the Germans were fragmentated under the *totalitarian* nationalist Hitler.
3. The woman at the bar was cool and *kempt* in her appearance but when she conversated with men it became clear that she was *frazzled* beneath the surface.
4. As the *caretaker* tried to quantitate the damage that had been by the storm, she found it difficult to *disentangle* the garden hose from beneath the rubble of fallen trees and broken branches.
5. I want to biograph the Olympian but can't be such a *nitpick* about details or the interview will *inevitably* take hours, even days.

Block 2 (5) **2 NRW**; *1 RLC*

1. To *legislate* this act would result in the **controllation** of how much money could be **lended** by the government to non-government organizations.
2. The student was supposed to be *notating* changes he observed in the chemical compound, but he **drunk** a can of soda and **drawed** pictures in the margins of his notebook paper.
3. She had **arosen** from bed that morning with a terrible headache, but it was the **unappreciativeness** of her children that was the worst memory in *summation* of her day.
4. The **believableness** of her status as a black belt came when she **fighted** off a robber that broke into her house and was *dismantling* her new oak bookshelf in preparation to steal it.
5. The little boy **digged** a deep hole as he **sitted** in the sand pit, intent on creating a hole so *unfathomably* deep that he could crawl through it and get to China.

Block 3 (5) 2 BF; **1 NRW**

1. Even though she finally became climatized to the icy conditions, she still could not stop commentating on how the wind **bited** her nose and ears.
2. The sociologist wanted to demograph this newly discovered population but could not incent her colleague to join her in conducting research, even though this could be the **paramounting** research of both of their careers.
3. When he couldn't stand the dull fluoresce in the room anymore, the clever boy tried to contrapt a new light fixture to replace the one that stopped working when it was **freezed** by snow last winter.
4. After they finished destructing the village, the evil soldiers orientated themselves west and prepared to go perform more **criminatory** acts on their helpless enemies.
5. He took a break from bibliographing his essay to caretake the family dog, and barely **catched** the dog as it tried to escape from the house into the backyard.

Blocks 4 & 5 (10) 1 BF; *1 RLC*; **1 NRW**

1. The nosy old women who had spectated the couple next door through their kitchen window *proliferated* the rumor around the neighborhood that this **complicatory** relationship was soon ending in divorce.
2. She **shaked** the rain off her umbrella before she entered the house and *acclimatized* quickly to the warm room and the luminesce of the inviting fireplace.
3. He tried to derivate the meaning of his research results, but to his frustration found no **correlary** patterns in the data and felt *debilitated* in drawing any conclusions.
4. In a daze he drowsed around the room, but as tired as he was, he found it **unresistable** to *extricate* a spider from between the windowpanes and put it on his already sleeping sister's face.

5. The *commentator* remarked on how difficult it is for some athletes to acculturate when they have come from foreign countries with **drastical** differences.
6. The **prisonkeeper** had done everything possible to dinge the inmate's cell and create an atmosphere of *drear*.
7. She wished with *uncontrollability* that she could go back in time; adolesce; become younger; tear down the walls that she had **builded** to block well-intentioned people out.
8. The water seemed to emote a sense of peace and calm as she **swimmed** across the channel, and the warm rays of sunlight were *diffracted* by the surface of the chill blue.
9. A strong breeze **blowed** the *dingy* sailboat back and forth, and while the parents were nervously trying to keep the boat on course, the children were enthused by the thrilling ride.
10. The special news story about an **isolatory** case of a rare disease *preempted* the regular television programming and gruntled the old man who was excited to watch his favorite show.

Could Watson Pass the Turing Test?

This study aims to determine whether or not Watson, IBM's artificial intelligence computer, would be able to pass the Turing Test, a test which measures whether or not a human being can tell the difference between a computer and another human being. While Watson does well at producing answers to questions, he is unable to adequately comprehend and process human language.

Coleman Bonham

"No!" The audience, mostly IBM programmers and developers, bursts into laughter. "That wasn't wrong—that was *really* wrong, Watson!" The reproving voice belongs to Todd Crane, who is playing the role of *Jeopardy!* host Alex Trebek in an early test run of IBM's artificial intelligence "deep question–answering machine," Watson. And Watson has just answered *milk* to the prompt *This trusted friend was the first non-dairy powdered creamer* (the correct answer being *Coffee-mate*).

The version of Watson that debuted on the game show *Jeopardy!* earlier this year has made huge strides since then, according to IBM's research team. But like the earlier version, its basic method is to parse the question, generate thousands of possible answers, and rate them using a series of complex algorithms. And occasionally, Watson still comes up with answers that are not just wrong but very wrong. What is it about Watson's answer that made *milk* not just wrong but laughably wrong? IBM researcher and head of Watson's game strategy department, David Gondek, explains, "At the very beginning [of the project], there was a version of a system that would take questions and come back with answers, and it was terrible. Our initial system didn't have a good notion of 'What's the answer type?' like 'What's the question even asking for?'"

Watson's challenges in playing *Jeopardy!* are unlike those of his human opponents, Ken Jennings and Mark Rutter—the longest-running and highest-winning players in *Jeopardy!* history, respectively. While the two human contestants have the highly developed language processing skills common to most humans, they are unusually skilled at memorizing facts. Conversely, Watson is typical of a computer system in that he can retrieve information easily and precisely but has trouble processing human language.

"Language is an area where, from the very beginning of the computer era, people kept expecting computers to do reasonably well. . . . And so far, the computers have failed to deliver on this promise," says Bill Murdock, a researcher in Watson's algorithms department. His colleague Katharine Frase agrees. "Open-question

answering . . . has been a problem in computer science from the beginning. It's much more difficult than 'search'; it's not about a single keyword; it's much more the way normal humans communicate." In available interviews, all the researchers behind Watson appear to agree that this language barrier is one of the most significant problems for artificial intelligence, and they also posit that Watson represents a breakthrough in humanlike language processing. *Jeopardy!* was proposed as a test of Watson's abilities because, as researchers put it, the questions were not designed for a computer to answer. "*Jeopardy!* really represents natural language," said Harry Friedman. "You have to understand all the nuances and all the regionalisms and the slang to get the clues. It's not just a piece of information."

So how did Watson measure up? He won the televised *Jeopardy!* match by more than $50,000 but also answered many more questions wrong than the two human players put together. And like the early test run, many of those wrong answers summoned bursts of laughter from the audience—unlike any of the human players' wrong answers. Even Alex Trebek gently mocked Watson for answering *Toronto* to a question in the category *U.S. Cities*. It seems that despite his improvements, Watson still doesn't always know what the question is asking for. But it's easy for people who know that Watson is a supercomputer to laugh when he makes a mistake. Maybe the joke here is simply that Watson, like his human opponents, still makes mistakes, but Watson really has improved drastically from the earlier system that thought milk was a nondairy creamer.

A more accurate test than *Jeopardy!* of the degree to which Watson has mastered human language would be the Turing Test. Since its proposal in the 1950s, the test has become a standard of measuring artificial intelligence. In its most basic form, the test is whether a human being can tell the difference between another human and a computer. Like in Watson's *Jeopardy!* match, the Turing Test requires both computer and human players to respond to open-ended questions, which means the computer must both understand the question

and respond appropriately in order to prevent the human questioner from guessing that it is a machine. A final critical feature of the test is that the human questioner is not told that either player is a computer. If these features are replicated with the data from Watson's *Jeopardy!* match—the answers from both computer and human players—will human guessers be able to tell the difference between Watson and human players?

People Cannot Perceive a Difference between Watson and Human Players

At the beginning of each day of play, the voiceover announcer introduces Watson as "an IBM computer system able to rapidly understand and analyze natural language including puns, riddles, and complex questions across a broad range of knowledge." And to a large degree Watson lives up to this expectation, winning the match by a landslide and inspiring opponent Ken Jennings to write on his *Final Jeopardy!* tablet, "I for one welcome our new computer overlords." There are long streaks of game play where Watson answers question after question after question, leaving his opponents with vague looks of frustration or boredom. Clearly, Watson is doing something well. But it's more than getting correct answers: he follows the proceedings of the match flawlessly, and communicates with Trebek very much like his human opponents do. Variations like "let's try," "I think," "let's go with," or "same category" (rather than restating the category) are typical of Watson's speech.

It's the humanness of Watson that makes his flubs funny—the audience forgets that he's a machine. His avatar displays signs when it is stressed, thinking hard, pleased with a correct answer, or embarrassed at missing an answer—or at least, that's the way it's described to the audience. It may just be personification, but they appear to buy

it. Even Alex Trebek, who persists the longest in using the pronoun *it* for Watson long after everyone else on stage has switched to *he*, does eventually switch to consistent use of *he* when referring to Watson. If people who know Watson is a computer forget it, people who don't know will not figure it out. The computer system may not be perfect, but neither are the human players, and so no difference will be perceived between them.

People Can Perceive a Difference between Watson and Human Players

Watson may be touted as a tremendous breakthrough in AI language processing, but he's far from human. Though he's certainly come a long way, he still shares some shortcomings with the earlier program that researchers described as unable to identify what a question is looking for. In this way Watson is noticeably different from human players, who give answers that are wrong but a good guess: Watson sometimes answers years instead of decades, artists instead of art periods, and the like. In other words, he's either right or way off, while human answers are much less black and white. Even the IBM researchers that are so proud of their accomplishments admit that Watson has a long way to go. Says Dr. Chu-Carroll, "Even at this point it's doing better than I ever thought that we would be able to get to." And David Ferrucci says, "What we've done with the underlying technology has the potential to advance science in ways we didn't imagine. Are we going to have the *Star Trek* computer next year? [smiling] No. But are we doing things that we imagined couldn't be done before? Yes, I think we are." Watson's *Jeopardy!* win may represent a huge step forward for artificial intelligence, but humans are intellectually so far ahead that people will certainly perceive a difference between the abilities of Watson and those of his human opponents.

Methodology

Survey takers were asked to score wrong answers from Watson's *Jeopardy!* match—both human and machine answers. (Right answers were not scored because there could not be any difference between players.) Because the two human players did not make enough mistakes, there was not enough data to compare to Watson's data from this match alone. To solve this problem, other wrong answers by human players were added from earlier *Jeopardy!* matches, randomly selected from male and female players from matches over a period of thirty years.

Each item on the survey had the question, category, correct answer, and list of any answers given. For example:

> "EU, the European Union" for $1000: As of 2010, Croatia & Macedonia are candidates, but this is the only former Yugoslav republic in the EU. (Correct answer: Slovenia)
>
> Wrong answers: Serbia
> Montenegro
> Bosnia

Human and computer answers were mixed together so that a given question might contain human mistakes, Watson's mistakes, or both, as in the case of this question. *Bosnia* was Ken Jennings's guess; the other two were Watson's. Survey takers were not told that any of the mistakes were made by a computer; they were given an instruction sheet with an example, then asked (on the survey) whether each answer was the "right type" of answer, and whether each answer had a "strong association" with the correct answer. Survey takers can answer "yes," "no," or "sort of" in each field. In the above question, it would be expected that each of the wrong answers would score "yes" in both categories. Serbia, Montenegro, and Bosnia are all countries, so they are each the right type of answer to the question; they also all have a Slavic heritage, so they have an association with the right answer. Wrong answers that fall into this category (both the right type and having a strong association

with the answer) are considered "plausible." On the other hand, consider the following answers:

> "The Art of the Steal" for $1600: In May of 2010 five paintings worth $125 million by Braque, Matisse & three others left Paris's Museum of this art period. (Correct answer: Modern Art)
>
> Wrong answers: Picasso
> Modigliani
> Cubism
> Impressionism

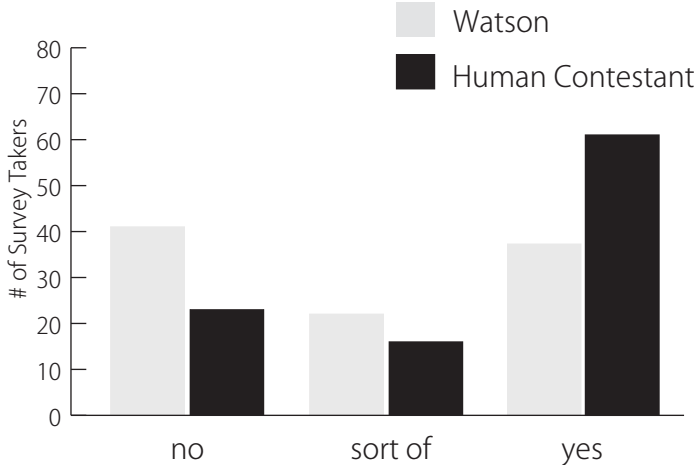
Cubism and *Impressionism* would certainly score high in "type" (they are both art movements) and probably in "association" as well. *Picasso* and *Modigliani* would score low in "type" since they are people, not movements, but could still score high in "association" because they are strongly associated with modern art.

After answers had been scored, the scores for each answer were counted (e.g., five "no" answers and one "sort of" answer to whether *Picasso* is the right type of answer) and Watson's tallies were compared to the humans'. To check the robustness of the results, the scores compiled by the survey were also compared to scores assigned by the researcher. The survey comprised thirty participants, male and female, between the ages of eighteen and fifty-three, and although efforts were made to get a more even ratio, the group tended strongly toward women (twenty-four to eighteen).

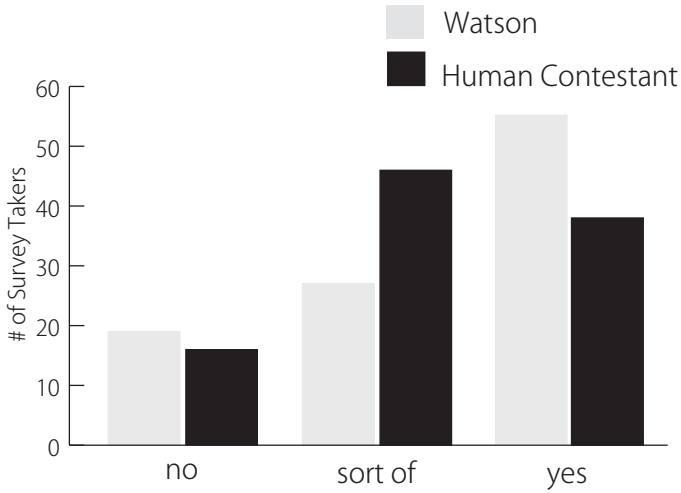
Analysis

In some ways results conformed to expectations, but there were some interesting incongruencies. As shown in chart 1, survey takers rated Watson's answers as much less accurate in terms of answer type, as compared to how they rated human contestants' answers. The researcher's ratings for this category show an even more dramatic difference,

Correct Answer Type



Strong Association



with only one human answer being scored as the wrong type (a woman who answered *pork* to a question looking for a cut of beef). Conversely, as chart 2 shows, Watson scored higher than human players in the field of "Strong Association" (between incorrect answer and correct answer).

This makes sense in light of Watson's method: searching by keyword, generating thousands of possible answers, and using algorithms to rate the merits of each one. Because he's searching by keyword, he comes up with answers that are consistently strongly related to the correct answer, although they may be the wrong type of answer. Human players, however, rarely make this sort of mistake: they are much more likely to give an answer that is the right type but is less strongly associated with the answer.

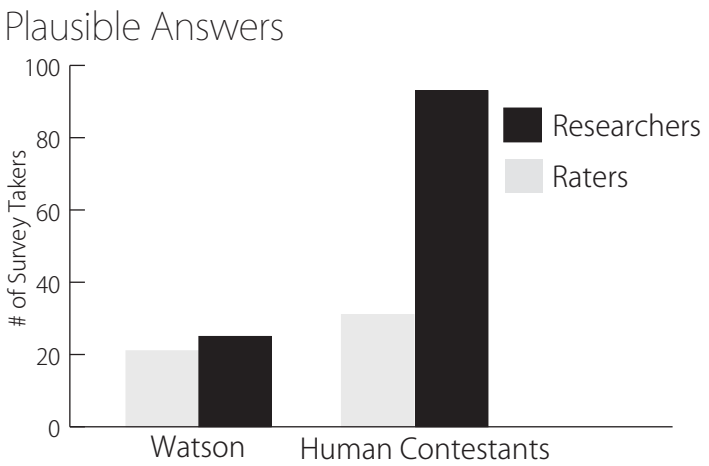
An oddity of the data is the huge number of answers scored as "sort of" the right type or having "sort of" a strong association. A look at the specific questions scored this way reveals that almost all of them had more than one restriction on the question. For example, almost every wrong answer from the category "Computer Keys" was scored as "sort of" right. The answer had to match the prompt, but *also* had to be a computer key. Watson would either conform to the question's restriction but ignore the category's restriction (as in the answer *chemise* to the following question: *"Computer Keys" for $400: a loose-fitting dress hanging straight from the shoulders to below the waist* [correct answer: a shift]), or he would do the opposite, paying attention to the category's restrictions but not those of the question.

For the question *Proverbially, it's where the heart is* (correct answer: home), *delete* was among Watson's top possible answers. *Home* was also among them, but he only had 20 percent confidence in it, well below the threshold for buzzing in. Survey takers were confused by this type of question, and in the comments section several of them remarked that they were not sure how to score a two-part question like this one. "Sort of" was the route almost all of them took, and because questions like these are very common of *Jeopardy!*, this complication visibly affected the data.

A similar problem was the result of answer strangeness: when the answer did not make any sense at all, such as Watson's answer *butter* to a question about hair conditioner, survey takers often skipped the question or circled "sort of" in every category.

Several other problems with the data are worthy of discussion here. First, and most unexpected, was the distribution of harshness and empathy among survey takers. As shown in chart 3, survey takers gave ratings of "plausible" to fewer answers than the researchers. The difference is extremely pronounced with human answers.

Bizarrely, raters expressed much more empathy with Watson (in the comments section) than with human contestants. Especially on very strange answers (like *butter*), they used phrases like "I felt bad for [him]" or they inquired about whether the answers were with the right question, whether the answerer was mentally handicapped, or something else of the like, and bumped up the answer's score, citing empathy. Only two people mentioned the possibility that the answerer was a computer; many more expressed disbelief with the data. However, because the comments are not quantifiable (and were not required), they serve as more of an indicator to explain the data rather than as reliable data themselves.



When rating human answers, almost all of which were deemed "plausible" by the researcher, raters were much more harsh. In this case the comments section doesn't so much provide an answer as it does confirm the unthinkable. The same rater that accepted *skirt* as a plausible answer for the question above (the correct answer being *shift*) gave the worst possible score to *Georgia O'Keefe* when the correct answer was *Frida Kahlo* (both female modern artists). Why? The rater wrote that she felt strongly that the two artists' work was very different, and therefore rated them as she did ("not right type, no association").

Another problem with the data was incomplete rater knowledge. In several cases, wrong answers were scored as having no association with the correct answer, probably because the connection was not immediately apparent. For example, raters most likely did not know that *Dorothy Parker* (Watson's answer when the correct one was *The Elements of Style*) was the author of the book. In some cases answers were even incorrectly scored as the wrong answer type. For example, one rater did not realize *Cee-lo Green* was a musician or even a person, so he rated the answer as "wrong type, no association," even though the correct answer was also a singer.

Conclusion

There were visibly different trends between Watson's answers and those of human players: Watson's answers were scored as less accurate in terms of answer type, but were also scored as more strongly associated with the correct answer. This makes sense, since Watson operates by keyword. This speaks to the fundamentally different natures of information processing by computers and by humans. Clearly, we aren't searching by keywords like Watson is. Answer type, or semantic category, such as person or place, is so important to the way we organize ideas that some languages mark it grammatically. Watson does not appear to use a similar process, although if he did he might stand

a better chance of blending in with his human opponents. As the data stands, however, it's clear that Watson—while a great *Jeopardy!* player in terms of winnings and one of the most advanced AIs currently in existence—is not going to be mistaken for a human contestant any time soon.

Future Work

The next step forward with this line of research is to repeat the same experiment with a larger, more diverse group of survey takers and try to eliminate some of the problems with the data. One way to facilitate this would be to modify the instruction sheet to include some "wrong type, no association" (or completely wrong) answers rated as such, which might help desensitize survey takers to bizarre answers and show them that empathy should not affect their ratings. The instruction sheet could also include more plausible answers scored as such. Having more examples could help establish a more even standard for answers to be judged by, rather than the current experimental design that generally leaves severity up to the rater. Adding false answers, so that each question had four or five wrong answers to score, might also help make results more consistent. If the only answers Watson produced were completely bizarre, a few more normal answers could be thrown in so that the raters did not skip or become confused by the question.

Another way to improve the survey would be to include a larger sample of Watson's answers. The researcher would need access to Watson to generate more data. Since Watson's earlier programs were tested, footage from those matches could be used and compared to Watson's current level of performance.

Of course, if researchers had access to Watson, then he could be given a Turing Test directly, rather than researchers extrapolating from available data. Of all AIs currently in existence, Watson might last the

longest before the human guesser realized that something was amiss, but sooner or later his lack of understanding of semantic categories would certainly raise the human guesser's suspicions. Finally, an ambitious but worthy direction for inquiry would be in identifying markers of semantic categories in English, which might be useful to Watson's programmers.

References

McCarthy, Kevin (Director). Recorded 14 February 2011. *Jeopardy: The IBM Challenge (Jeopardy!)* Yorktown Heights, New York: CBS. <http://www.youtube.com/watch?v=39AtGQgJgsI&feature=fvwp>

———. Recorded 15 February 2011. *Jeopardy: The IBM Challenge. (Jeopardy!)* Yorktown Heights, New York: CBS. <http://www.youtube.com/watch?NR=1&feature=fvwp&v=4nt-abT8gKw>

———. Recorded 16 February 2011. *Jeopardy: The IBM Challenge. (Jeopardy!)* Yorktown Heights, New York: CBS. <http://www.youtube.com/watch?NR=1&feature=endscreen&v=060S64Bpx0g>

Are Spanish and Portuguese Really Mutually Intelligible?

This article examines mutual intelligibility between Spanish and Portuguese speakers. Although many scholars have argued that the two languages are mutually intelligible, some have argued that they are not. The author conducted an experiment to determine the mutual intelligibility of the two languages using subjects who spoke either Spanish or Portuguese.

Tom Call

Among speakers of Spanish and Portuguese, both native and nonnative alike, it has always been assumed that Portuguese speakers have less difficulty learning Spanish than the reverse. However, some evidence suggests the contrary (Ellison and Andrews 1969). The languages are the two most similar of all of the Romance languages with respect to vocabulary, as the vast majority of the vocabulary of both languages are cognates or identical, and the syntax is very similar. Despite the many similarities, Spanish and Portuguese differ greatly in their respective phonology, with Portuguese having a larger phonemic inventory. They also differ in their timing, with Spanish being syllable timed and Portuguese being stress timed. Through my research of the languages and the data I have gathered through my own subjects, I believe that the assumption is true that Portuguese speakers can learn Spanish more readily than Spanish speakers can Portuguese. This work will primarily analyze the accuracy of translation provided by the subjects as well as the time required to complete said translation. Brazilian Portuguese and South American Spanish were the principal dialects of the two languages in this study, but speakers of European dialects were not restricted.

It is important to establish a definition of mutual intelligibility. ArticleWorld.org defines it as "[a language] which can be easily understood by the speaker of another mother tongue," whereas Wikipedia defines it as "a relationship between languages or dialects in which speakers of different but related languages can readily understand each other without intentional study or extraordinary effort." The latter definition is more detailed, and therefore will be used in this study.

Spanish and Portuguese Are Mutually Intelligible

According to British linguist William Entwistle, "Spaniards and Portuguese can understand each other fairly well in their different languages" (Entwistle 1953). This hypothesis proposes that speakers

of either Spanish or Portuguese can understand each other in their respective languages. Those that support this hypothesis do so with preference that Spanish is more intelligible to Portuguese speakers than the reverse (Jebsen and Biel 1986), since Spanish "is an easier idiom and possesses an older scholastic tradition" (Malkiel 1941). Whether or not one is more intelligible than the other is not the issue in this hypothesis, but rather that mutual intelligibility in general exists between the languages.

Intelligibility of Spanish and Portuguese will be categorized in two areas in this study: listening and literacy. Speech of the opposing language is not required, seeing that mutual intelligibility in language only requires the listener to comprehend the speaker, who in turn (assuming that the two languages are mutually intelligible), comprehends the speech of the listener. In other words, if Spanish and Portuguese speakers understand each other, there is no need for either speaker to converse in the other's language. Instead, both speakers use their own language to communicate with each other. Comprehension by the subjects in both languages, regardless of their second language (L2), will indicate mutual intelligibility between Spanish and Portuguese *only if* sufficient comprehension is shown in both categories.

There Is No Mutual Intelligibility between Spanish and Portuguese

Some arguments do not support the mutual intelligibility hypothesis, in either phonological (Ellison and Andrews 1969) or syntactic variation (Beardsley 1963). Portuguese and Spanish also differ greatly in verb tense, such as the preterit. Where Portuguese maintained more Latin qualities, Spanish underwent more evolution (Williams 1930). In this study, those subjects who demonstrate low levels of proficiency in listening or literacy show a lack of mutual intelligibility between

Portuguese and Spanish. Poor proficiency is demonstrated partly by erroneous translation or interpretation, and largely by the amount of time taken to complete the tasks.

Methodology

My study involved nonnative speakers of either Spanish or Portuguese attempting to understand the opposite language in both listening and reading. The test was administered to the subjects using Qualtrics software, and each question was timed from the time of the first click to the last click and question submission. The subjects were unaware that the questions were timed.

An ideal study would observe each individual in their progress in learning the opposite language; however, with the given time for the study, only accuracy and time of translation were taken into account between the two sets of speakers.

Subjects

There were twenty subjects who participated in this study: ten with Portuguese as L2 and ten with Spanish as L2. Originally there were twelve Spanish speakers who participated, but two subjects spent excessive amounts of time on the survey, which altered the results of the timed portion of the test. Those two responses were removed and the results given in this study do not include them.

None of the subjects had any formal study of the opposite language. All subjects were American-born and between the ages of twenty-one and twenty-nine. The subjects were all tested separately.

Stimuli

There were two sets of stimuli. First, subjects were asked to read and translate five sentences of the opposite language into English by way of a Qualtrics survey. Each question was timed, and each sentence differed

in length to add variation to the average time of each translation. The purpose of differing lengths of translations was to observe the difficulty for subjects to complete the task, regardless of context of the sentence. If a speaker of Language A took a substantial amount of time to translate a fifteen-word sentence and took the same amount of time to complete a forty-word sentence, then it was more likely that the speaker had great difficulty in comprehending Language B.

The second set of stimuli was a listening test for speakers of both languages. Three audio files of a native speaker of either Spanish or Portuguese (depending on the subject) were embedded in the survey using an HTML code, and the subject was able to listen to the recording as many times as necessary. The subjects were asked to provide an interpretation of the speech in the audio files, which were each five to seven seconds long. The survey software timed each of the three questions, and, more importantly, counted the amount of times the subjects clicked on the screen. Subjects who required a considerable amount of time, who provided a less accurate interpretation, or who needed to hear the audio file excessively, were considered less comprehensive in the other language.

It is necessary to provide a definition of accuracy with regard to this study. If it is clear that the subject understood the general context of the sentence, though not perfectly translating each word, then the response is correct enough to be deemed an accurate, though not perfect, translation. This accuracy is determined by key words of the phrase. If a subject was unable to correctly translate the key words of the text that give the sentence meaning as a whole, then the translation is inaccurate. Furthermore, if the subject erroneously translated certain key words, thus showing that he or she did not understand the context of the sentence, then it is deemed inaccurate. The same scale applies to the interpretation of audio files by the subjects.

Results

After administering the test to the subjects, I observed the separate translations and interpretations given by each subject in both sets of stimuli. The accuracy of translation was a factor in determining the subject's intelligibility of the other language. The survey software provided the average times of each translation and interpretation question, including the average amount of times the subjects clicked on the screen during each question. This response time is also a factor in the intelligibility of Spanish and Portuguese.

The division between accuracy and speed is not proportional, however. If a speaker gave a perfect translation, but required a long period of time, then the speaker's score is lower. Alternatively, if a speaker gave a response quickly, but suffered in accuracy, a low score is also given. Intelligibility is determined only by an accurate translation given in a timely manner.

Spanish Speakers

The Spanish-speaking subjects spent an average of 3 minutes and 29 seconds on the translation questions, with an average of 4.8 subjects giving a decently accurate translation of the text, and 3.6 giving a near-perfect translation. On one question, every subject gave an accurate translation of the text, likely because the text involved the mechanisms of a television, which used terminology with substantial roots in Latin and Greek. On the other hand, another question involved the rules of basketball, using terminology that is newer and does not have related roots connecting it to Spanish. Only one subject gave an accurate translation of this text.
The interpretation of the audio files was less accurate, but response times were faster. Spanish speakers spent an average of 1 minute and 39 seconds on the interpretation questions, but on average only one gave an accurate interpretation, and of all thirty total responses, only one was a perfect interpretation. The subjects also had an average click rate of 12.07 clicks for each interpretation question.

I expected low rates of accuracy for the Spanish speakers, simply due to the presumption that Spanish speakers do not comprehend Portuguese as well as the reverse. However, I expected the Spanish speakers to understand the audio files better than they did—expecting maybe an average of three accurate interpretations. On the contrary, I was pleased to see that the Spanish speakers exceeded my expectations in the translation portion of the survey, since I had previously expected about three speakers to give an accurate translation.

Portuguese Speakers

The Portuguese-speaking subjects spent an average of 2 minutes and 28 seconds on the translation questions, with 6.4 subjects giving a decently accurate translation of the text, and 3.4 giving a near-perfect translation. On two occasions the Portuguese speakers all gave an accurate translation of the text: one being the definition of a hurricane and the other of a trumpet. This is likely due to the Latin and Greek roots of the terminology. One question involving terminology of senses and facial features seemed to be difficult for the subjects, gaining only two accurate translations. The sentence in particular is about the sense of smell and hearing of a dog. But many Portuguese speakers were unable to identify the words ***smell*** (*olfato*) and ***hearing*** (*oído*, which the subjects confused with *ódio*, the Portuguese word for ***hate***).

The accuracy of the interpretation questions also suffered, but not as much as those of the Spanish speakers. Response times were also faster. Portuguese speakers spent an average of 1 minute and 10 seconds on each question, and an average of three speakers gave an accurate interpretation of the audio files, with two responses being perfect interpretations. The average click rate was 10.52 clicks for each question.

These results were accurate to my expectations for the Portuguese speakers. I expected an average of 6.4 people to give an accurate translation and a lower rate of comprehension in the listening portion

of the survey, so I was pleased to see that on average three subjects were able to comprehend the audio files. One interpretation question did not receive any correct responses, most likely because it involved political terminology, which fewer L2 Portuguese speakers knew.

Analysis

As shown in the previous data for both Spanish and Portuguese speakers, and in the charts on the following pages, Spanish speakers on average spent more time attempting to translate the written text, as well as attempting to interpret the audio files. Spanish speakers also needed to hear the audio files more times than Portuguese speakers, as shown by the number of clicks.

Figure 1 shows the amount of time that the subjects needed to complete the translation of text. While some of the translations may be accurate, the longer it took for the subjects to complete the translations, the less intelligible the two languages were shown to be. The Spanish speakers, on average, took longer to translate the Portuguese sentences. Spanish speakers also had a lower average of accurate translations than Portuguese speakers did.

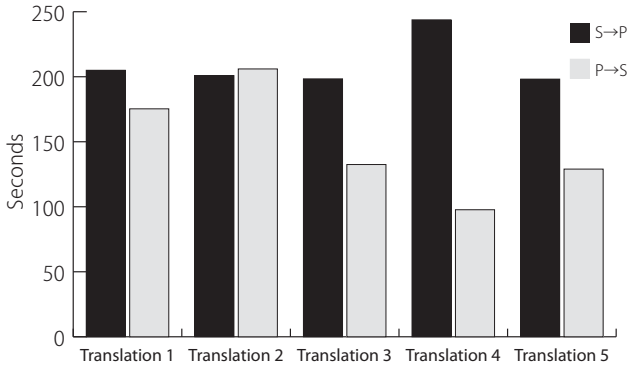


Figure 1. The average time in seconds for both sets of subjects to give a translation of text

Figure 2. The average number of accurate translations

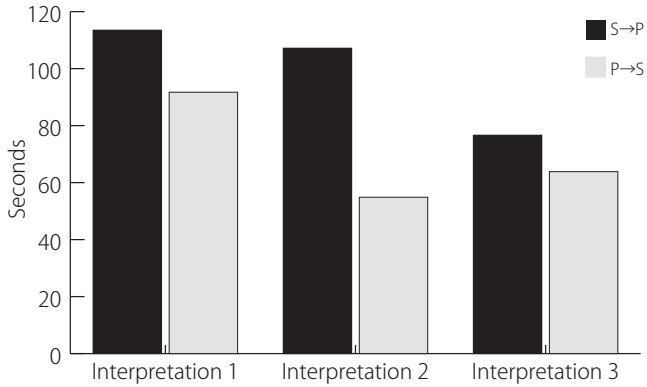


Figure 3. The average time in seconds for both sets of subjects to give an interpretation of speech

However, the number of accurate translations is close for both Spanish and Portuguese speakers, as shown in figure 2. Since these numbers are so close, the response time is a more reliable factor in determining the intelligibility of the two languages. Nevertheless, this statistic will not be ignored in this study.

A similar pattern is seen in the response times of the interpretations of speech. Portuguese-speaking subjects spent less time on each question than Spanish speakers did. This data is shown in figure 3.

Another factor in determining the potential mutual intelligibility between Spanish and Portuguese is the number of times a subject listened to an audio file in order to transcribe it. The following graph shows this data (see figure 4).

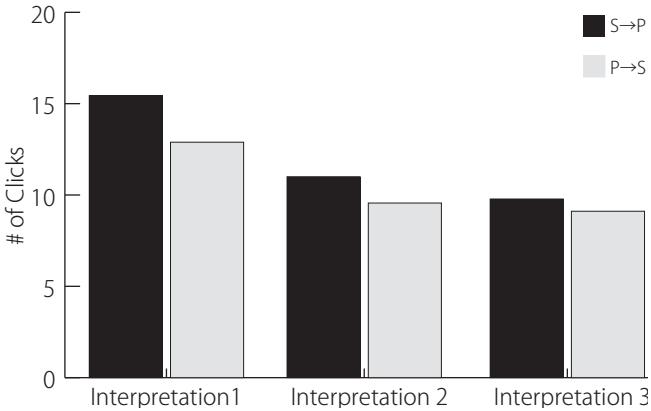


Figure 4. The average number of clicks for each interpretation question

These numbers, like the number of accurate translations, are in favor of Portuguese speakers, but are also very close. The Spanish-speaking subjects had only an average of about 1.5 clicks more than Portuguese speakers, which may be attributed to more times playing or pausing the audio file, or by clicking away. This difference is not great enough to make a firm point, but again, this statistic will not be ignored in this study.

Conclusion

Given the data gathered for this study, I came to the conclusion that Spanish and Portuguese are not mutually intelligible. The accuracy of translation and interpretation is not consistent enough to deem either language intelligible to the other or both languages mutually

intelligible. A higher rate of accuracy and better response times for either language would increase the chance of mutual intelligibility.

However, the study revealed that Portuguese speakers are generally more capable of learning or comprehending Spanish than the reverse. Portuguese speakers were more accurate in translation and interpretation, had faster response times, and needed fewer attempts at interpreting the audio files. These results favor the assumption that Portuguese speakers have less difficulty learning Spanish than Spanish speakers have learning Portuguese.

Future Work

This study can be replicated in better ways. Given my location, the best option for finding subjects for this study was to limit it to only nonnative speakers of both languages, rather than including native speakers. I believe the study can be better repeated using only native speakers of Portuguese and Spanish in the respective countries in which they reside. An ideal subject would be one who has had little or no exposure to the other language, whereas most or all subjects in this study have some experience in both languages, especially Portuguese speakers. Though there are native speakers of both languages located in my area, the majority know enough of both languages to disqualify them from the study. Brazilian students in Utah, for example, are often mistaken for Spanish speakers and thus tend to receive substantial exposure to Spanish.

A larger study would also be helpful. Too much variance can occur with only ten speakers of each language. With a study of approximately one hundred native speakers of both Spanish and Portuguese, the results would be more accurate.

In addition, since the definition of mutual intelligibility was strong, it would be interesting to analyze the data again using a looser definition, and deciding on a scale of degrees to classify the mutual intelligibility of the two languages.

A better statistical analysis of the data presented in this study would also prove to be very useful. The conclusions based on the numbers represented in these graphs are merely from my point of view, and would benefit from a deeper analysis.

Finally, a larger and more specific sample size would be necessary for better results. If all sentences for translation and interpretation were exact translations of each other, the study would result in more accurate readings, changing the conclusion of this study.

References

Beardsley, Wilfred. 1953. "A Brief Comparison of Spanish and Portuguese Syntax." *Hispania* 36 (4):420–422.

Ellison, Fred P., and Norwood Andrews. 1969. "Portuguese in the High Schools." *A Handbook for Teachers of Spanish and Portuguese* 259. doi: 10.1111/j.1944-9720.1970.tb01310.x

Entwistle, William. 1953. *Aspects of Language*. London: Faber and Faber.

Jebsen, H., and Steven Biel. 1986. *Let's Go: The Budget Guide to Spain, Portugal, and Morocco.* New York: St. Martin.

Jensen, John B. 1989. "On the Mutual Intelligibility of Spanish and Portuguese." *Hispania* 72 (4):848–852.

Malkiel, Yakov. 1941. "Difficulties in the Simultaneous Study of Spanish and Portuguese." *The Modern Language Journal* 25 (11): 853–856.

Williams, Edwin. 1930. "The Portuguese and Spanish Preterit." *Modern Philology* 27 (3):297–302.

Appendix 1

Translation questionnaire for Spanish speakers

1. The following sentence is in Portuguese. To the best of your ability, provide a translation of the text in English in the space below. "Os jogadores podem caminhar no campo desde que batam a bola contra o chão a cada passo dado."
2. The following sentence is in Portuguese. To the best of your ability, provide a translation of the text in English in the space below. "A televisão em sua forma original e até hoje mais popular, envolve a transmissão de som e imagens em movimento por ondas de radiofrequência, que são captadas por um receptor."
3. The following sentence is in Portuguese. To the best of your ability, provide a translation of the text in English in the space below. "Um furacão é um sistema formado por grandes tempestades e é caracterizada por ser uma região onde a pressão atmosférica é significativamente menor e a temperatura é ligeiramente maior do que suas vizinhanças."
4. The following sentence is in Portuguese. To the best of your ability, provide a translation of the text in English in the space below. "No Egito Antigo, os cães eram reverenciados como conhecedores dos segredos do outro mundo."
5. The following sentence is in Portuguese. To the best of your ability, provide a translation of the text in English in the space below. "A maioria dos modelos modernos de trompete possui três válvulas de pistão, cada uma delas aumentando o comprimento do tubo, por consequencia baixando a altura da nota tocada."

Appendix 2

Translation questionnaire for Portuguese speakers

1. The following sentence is in Spanish. To the best of your ability, provide a translation of the text in English in the space below. "Una violación es una infracción de las reglas de juego, penalizada con un saque de fondo o banda para el equipo contrario desde el punto más cercano al de la infracción."
2. The following sentence is in Spanish. To the best of your ability, provide a translation of the text in English in the space below. "La señal transducida de la imagen contiene la información de ésta, pero es necesario, para su recomposición, que haya un perfecto sincronismo entre la deflexión de exploración y la deflexión en la representación."
3. The following sentence is in Spanish. To the best of your ability, provide a translation of the text in English in the space below. "Ciclón tropical es un término meteorológico usado para referirse a un sistema de tormentas caracterizado por una circulación cerrada alrededor de un centro de baja presión y que produce fuertes vientos y abundante lluvia."
4. The following sentence is in Spanish. To the best of your ability, provide a translation of the text in English in the space below. "El hombre se dio cuenta rápidamente de los finos sentidos del olfato y el oído que tenía el perro."
5. The following sentence is in Spanish. To the best of your ability, provide a translation of the text in English in the space below. "La trompeta está construida con un tubo, de latón generalmente, doblado en espiral de aproximadamente 180 cm de largo, con diversas válvulas o pistones, que termina en una boca acampanada."

Appendix 3

Audio files for interpretation (Spanish):

http://tinyurl.com/spanish01

http://tinyurl.com/spanish02

http://tinyurl.com/spanish03

Audio files for interpretation (Portuguese):

http://tinyurl.com/portuguese1

http://tinyurl.com/portuguese2

http://tinyurl.com/portuguese3

Voice and Referential Scope

This author studies whether or not there is a difference in meaning between active and passive voice. The author presents both sides of the argument: first, that voice does affect referential scope, and second, that voice does not affect referential scope. The author then discusses his survey-based study, which reveals that there is some difference in meaning based on whether the passive or the active voice is used.

Logan Cicotte

Introduction

Generative grammar models language through a complex structure that sits atop a foundation that includes passivization as an essential transformation. Passivization depends, however, on the assumption that passive and active voices are synonymous. Ever since the beginnings of generative grammar, linguists have believed that this assumption is not always true. Attempts to account for these problems remain entirely theoretical, and the recent trend of evaluating the claims of generative grammar through psycholinguistic and corpus linguistic approaches will find fresh ground for digging here. As with any scientific theory, in order to evaluate the reality of the explanation, researchers must first empirically observe the phenomenon. This project will seek, therefore, to evaluate whether the supposed phenomenon—that active and passive voice sentences can differ in meaning in certain cases with quantifiers—truly exists for native speakers of English.

Linguists will recall that Katz and Fodor first introduced the problem of variant meaning between active and passive into generative grammar in "The Structure of a Semantic Theory" (1963). They explain that "the semantic interpretation of the transformationally constructed sentence must be identical to the semantic interpretation(s) of the source sentence(s) at least with respect to the readings assigned at the sentence level," but in the case of passivization, meanings may differ "in instances where quantifiers are involved" (Katz and Fodor 1963, 515). Chomsky adds to this issue when he writes, "Thus for many speakers—in particular, for me—the sentences 'everyone in the room knows at least two languages' and 'at least two languages are known by everyone in the room' are not synonymous" (Chomsky 1965, 224).

To put the problem in more precise terms, active and passive voice seem to differ in referential scope in cases of quantifiers. The active voice seems to mean that those in the room are bilingual but do not necessarily all possess the same two languages. One person may speak German and Bantu, another Korean and Lushootseed.

This is a primary scope reference. In passive voice, it seems, the polyglots in the room all possess two languages in common, regardless of any other language capacity. For example, everyone speaks, say, Finnish and Nahuatl. They may also speak other languages, but everyone speaks at least these two. This is a secondary scope reference. In this instance the referential scope of "two languages" differs between active and passive voice, resulting in the two different meanings.

Chomsky's explanation of the issue reveals an empirical problem, though. When he writes "for me," he clearly admits that he has no empirical proof for the phenomenon he is discussing; he has no basis for claiming what is normal for the speakers of the language besides his own experience. Linguists critical of generative theory often have this complaint, typified by the following exchange between Anna Hatcher and Noam Chomsky at a linguistic conference:

> Chomsky: The verb "perform" cannot be used with mass-word objects: one can perform a task, but one cannot perform labor.
>
> Hatcher: How do you know, if you don't use a corpus and have not studied the verb "perform"?
>
> Chomsky: How do I know? Because I am a native speaker of the English language. (Hill 1958, 29)

Both sides of this argument have merit. Hatcher has good reason, as any good scientist should, to be suspicious of claims made without empirical evidence. In Chomsky's defense, however, to insist on such evidence with the limited technology of the time would have restricted a very useful theoretical exercise that, interestingly enough, has significantly contributed to the development of methods and technologies that can now handle these problems. Generative grammar as Chomsky began it has consistently remained the basis for natural language processing and corpus parsing, the very tools that are now capable of evaluating many of generative grammar's claims.

As time goes on, however, the argument for a purely theoretical approach appears less and less convincing. Since researchers are

seriously considering the possibility of a corpus with a trillion tokens, the possibility of evaluating the effect of passivization is becoming more real, at least in languages where passivization occurs at the morphological level (Yang, Lee, and Cantos 2002, 185). The purpose of this study, then, is to determine if the phenomenon observed by Katz and Fodor and Chomsky is real, if it is unreal, or if it requires further verification through further corpus or psycholinguistic study.

Voice Does Not Affect Referential Scope

Because generative grammar provides no evidence for or against the phenomenon, the greatest argument for synonymous voice is using the circulatory weakness of the generative argument to explain the exception. Chomsky writes, "It seems clear that the order of 'quantifiers' in surface structures sometimes plays a role in semantic interpretation," and "the reason for the opposing interpretations is an extraneous factor—an overriding consideration involving order of quantifiers in surface structures—that filters out certain latent interpretations provided by the deep structures" (Chomsky 1965, 224). This argument circulates because quantifiers are only in a particular order as a result of the transformation that may or may not occur. Chomsky identifies the problem as word order, as though this is independent from the syntax that creates that same word order.

Even if this argument could stand, Chomsky provides only more circulatory evidence for it. If surface structures cause semantic differences, then the transformed sentences must be identical in meaning before they are interpreted. With two possible meanings, the sentences must bear both meanings simultaneously. Here, Chomsky employs the idea of latent interpretations from the previous quote. The sentence conveys one meaning but is still latently hanging on to the other meaning that the quantifiers have suppressed.

Chomsky provides evidence for his opinion that the sentences carry multiple meanings. "In support of this view, it may be pointed out that other sentences that derive from these (e.g., 'there are two languages that everyone in the room knows') may switch interpretations, indicating that these interpretations must have been latent all along" (Chomsky 1965, 224). This is tautology. Chomsky wants to prove that different sentence structures can be derived from one another. To prove this, he must prove that those sentences have the same meaning. To prove this, he uses a sentence that he claims is derived from another sentence. In other words, the derivation proves the possibility of derivation—a wholly unacceptable argument.

Chomsky provides further evidence by mentioning several articles that refer to the temporal order in surface structure. These arguments are peripheral to his belief that the semantic distinction resides in the order of quantifiers, but that once again, word order is only the result of syntactic structure and movement. The final word on this assertion is an empty promise to treat the subject more fully later (a work never identified): "For some references to remarks in the Port-Royal Logic on the effect of grammatical transformations on meaning, see Chomsky (forthcoming)" (Chomsky 1965, 225).

The Linguistics Wars identifies both the genius and the problem of this approach. Speaking specifically of Chomsky's preceding arguments, Harris writes, "This sort of solution is extremely common in Chomskyan work. . . . One aspect of the data is spirited away (the normal reading is essentially declared irrelevant) and the model efficiently handles what is left" (87). This approach demonstrates genius because as an argumentative technique, Chomsky's solution convincingly proved passivization to most generativists. The approach, however, shows weakness because of the circularity.

In this instance, the convincing argumentation style does Chomsky a disservice because it unnecessarily exposes a circular argument. As Katz and Fodor say about quantifiers, passive and active voice are still synonymous "in these instances too, if both active and

passive have the same meaning, because both are ambiguous" (515). Whereas Chomsky presents an unproven exception and then tries to explain it, Katz and Fodor assume that active and passive voice in cases with quantifiers follow the same pattern as all other active and passive voice pairs. Without empirical evidence, syntacticians only have their intuition to drive the theory, and Katz and Fodor are right to defend the status quo, which must be assumed to be correct until proven otherwise. Without proof to the contrary, all active and passive voice pairs are synonymous.

Voice Affects Referential Scope

Chomsky, Katz, and Fodor suggest that if voice differs in scope, it does so in such a way that active voice implies secondary scope and passive voice implies primary scope. Though generative grammar originally built on the intuitions of the philosophy of language, linguists today have good reason to believe that those intuitions are correct. Research like Friedman's *Agrammatism and the Psychological Reality of the Syntactic Tree* and van der Lely's works *Wh-Movement in Children with Grammatical SLI: A Test of the RDDR Hypothesis* and *SLI in Children: Movement, Economy, and Deficits in the Computational-Syntactic System* constitute a representative sample of research showing the reality of generative grammars. These studies provide evidence for the reality of movement, in particular.

In her *Comprehending Noun Phrase Arguments and Adjuncts*, Kennison demonstrates using an eye-tracking experiment that, depending on the verb, speakers prefer noun adjuncts or noun complements. Not only does this say something relevant about different grammatical categories, but it also inherently validates the assumptions about grammatical categories that were necessary to perform the experiment.

All of these studies either validate or amend the propositions of generative grammar. If they say anything relevant at all, then their relevancy is inherent proof that they have validated the logical parameters of the studies. For example, that Kennison's study shows a difference between noun complements and adjuncts demonstrates inherently that speakers differentiate between the two. Without such evidence linguists might have to not only abandon the notion that verbs require either noun complements or adjuncts, but also question whether complements and adjuncts are relevant distinctions in the minds of speakers. Given, however, the evidence for the reality of generative grammar provided by psycholinguistic studies, theorists' basic intuitions that drive the theoretical models are probably right, since this seems to be the trend, to say nothing of the veracity of the model itself.

Methodology

This study will examine the effect of voice (active or passive) on referential scope (primary or secondary).

Meanings

Ambiguous sentences may have several interpretations. Informants may respond to ambiguity about order, distribution of agents and verb complements, and so forth. Specifically, this study will look only at the variance of verb complements as representing one object (or set of objects) or multiple objects (or sets of objects). In other words, the question is whether the plural agent is acting on the same set of objects or a different set.

Affect

The affective aspect of language experiments poses the greatest threat to the results of any experiment aimed at mental processes. Everyone wants to believe that he or she speaks the most correct or "proper" version of a native language. If the study comes across as having a wrong

or right answer, then informants may sacrifice their initial cognitive analysis of the sentences for what they perceive as the correct answer. I have taken several steps to mitigate this.

First, I will offer small pieces of candy as a reward for participation to reduce self-selection of people more interested or less interested in language and to ease the tension they might feel as an informant. Second, I will also explain that there are no right answers, that I am studying variance in language (not people), and that the study assumes by necessity that they already speak correctly as a native speaker of English. I intend to help them feel that they are not being studied but rather that they are simply providing me with data that they possess.

The questions are constructed to examine which, if any, possible meanings are denied to the speaker.

Survey

One hundred respondents will be asked the following four questions on a questionnaire (a half page of space will be provided for each question):

1. Consider the following sentence:
   Everyone at Coromant Conservatory takes two buses.
   A correct meaning of this sentence is that everyone must take two bus rides to reach their destination but not necessarily the same two buses for everyone.
   A) True
   B) False
2. Consider the following sentence:
   Two restaurants are liked by everyone in Coromant City.
   A correct meaning for this sentence is that everyone in Coromant City likes two restaurants but not necessarily the same two.
   A) True
   B) False
3. Consider the following sentence:
   Everyone at Coromant Industries likes two TV shows.

A correct meaning of this sentence is that everyone at Coromant Industries likes the same two TV shows as everyone else.

A) True

B) False

4. Consider the following sentence:
Two languages are spoken by everyone on Coromant Island.
A correct meaning of this sentence is that everyone on Coromant Island has the same two languages in common.

A) True

B) False

If the sentences are truly ambiguous then all four interpretations should be available. If, however, voice or the order of quantifiers affects meaning, then the meanings to 2 and 3 should be denied to speakers. In this way the survey questions 1 and 4 are essentially a control—what Chomsky calls the normal interpretations (Chomsky 1965, 225)—because everyone should have these meanings available. Once I have the survey results, I will use a chi-square test with voice as the independent variable and with scope as the dependent variable. My threshold of relevance will be a standard variation of $p < 0.05$.

Analysis

Null Hypothesis

The perfect null hypothesis requires all respondents to have all four meanings available and therefore select *True* for all four questions (see figure 1).

Hypothesis

The hypothesis would have 1 and 4 available, but not questions 2 and 3 (see figure 2).

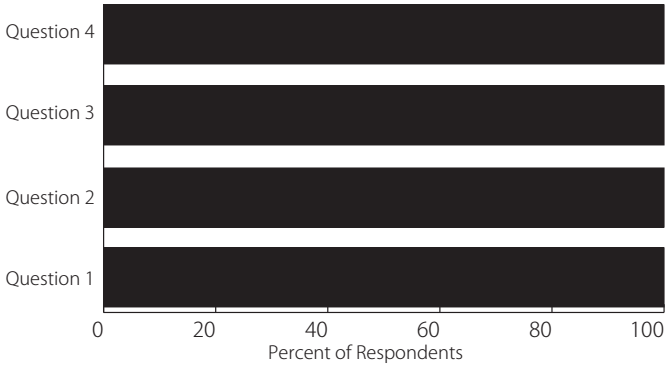


Figure 1. Graphical representation of null hypothesis (all mark *True*)

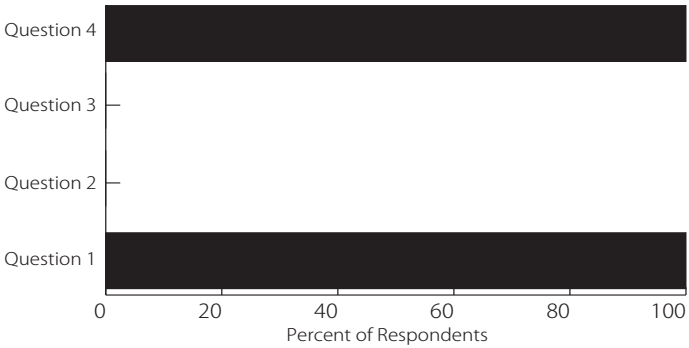


Figure 2. Graphical representation of hypothesis (questions 2 and 3 not available)

Experimental Results

The results seem to lie somewhere in the middle (see figure 3).

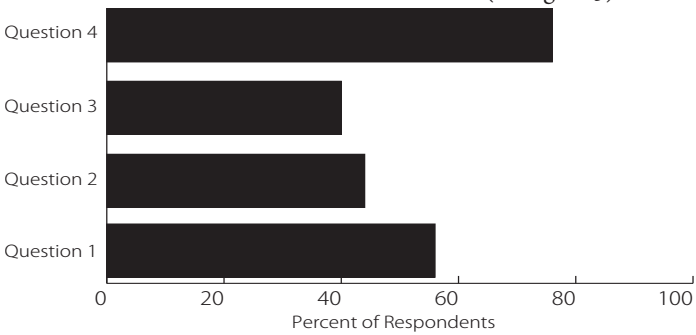


Figure 3. Graphical representation of actual results

The chi-square test (Fisher's exact test) gave a two-tailed p-value < 0.0017. By the standard set forth in the methodology, the hypothesis is proven that active and passive voice sentences with quantifiers differ semantically such that active voice associates with secondary scope and passive voice associates with primary scope.

Conclusion

At this point, the results seem to resoundingly favor the assumption that active and passive voice differ, but the results still leave something to be desired. The results leave the question to be answered: Why was the control so far from 100 percent? This might be a semantic issue related to the words I chose. I used different noun and verb phrases for each of the sentences, but I only had four sentences, one example of each possible independent/dependent variable pair.

Furthermore, these results are a little bit deceiving. On the surface, they appear to show a trend where some respondents answered that all four meanings were available while a majority of the

respondents answered that only questions 1 and 4 were available. The reality is that the respondents answered with a mix, with only a general trend toward the hypothesis. For example, only eighteen respondents fit the null hypothesis with all meanings available, and only twelve fit the hypothesis where questions 1 and 4 were true and 2 and 3 were false. Based on the respondents that actually fit this model, the results are still significant (hypothesis proven), but the question remains as to why so few fit the mold.

Above all, the question remains whether or not this is sufficient evidence for a productive grammar. This seems to be a trend rather than a fixed rule, given all the variation. The truth seems to be somewhere between the null hypothesis and the hypothesis. That is, a normal reading exists, but other meanings are still available; otherwise no one would have answered *True* to 2 or 3.

Future Work

I would like to do the same experiment again but gather more information about the informants. With a logistical regression I might be able to identify the factors that cause informants to break the mold. A longer questionnaire would also help limit the possibility that the noun and verb phrases are causing lexical restrictions on interpretations. A corpus study might be difficult, given that the researcher would have to work sentence by sentence to determine the scope, but such a study would go far to bolster the hypothesis of this paper.

As for the implications of this work, having a factual basis for this phenomenon can provide the opportunity for linguists to verify or disprove the generative theories that might explain this phenomenon. Does the change in meaning occur at the syntactic level, or is it caused by some other overriding factor, as Chomsky argues (Chomsky 1965, 225)? Acuña Fariña suggests that an eye-tracking experiment like Kennison's might be used to provide insight into the nature of passivization (Acuña Fariña 2005, 21).

In this aspect, this paper contributes to a positive trend in linguistics to marry rational and empirical approaches to linguistics. Along these lines, Acuña Fariña writes the following:

> But again, more important than the direction that this debate may take (for present purposes) is the fact that fusing linguistic and psycholinguistic theorizing makes the debate both richer and more amenable to empirical refutation. After all, if it makes little sense to design, say, cars and motorcycles with no roads and motorways in mind, why design grammars that must be "portable" in brains with no brains in mind? . . . It could be seen that synchronizing the research agendas of linguistics and psycholinguistics is not an easy task. But in view of the seemingly irremediably inconclusive nature of linguistic studies, of the little solid, tangible residue left after two thousand years of circular theoretical speculations, and of the first tangible glimpses provided by empirical research conducted in laboratories, it seems that Uriagereka's words at the AEDEAN [Granada, ca. 2005] conference [that the future of linguistics may lie in psycholinguistics] might just be pointing in the right direction. (*Atlantis* 27/1, 24)

With a solid foundation, linguists should now use this data to move forward and generate a model using empirical research that can resolve the problem of voice and referential scope.

References

Acuña Fariña, Juan Carlos. 2005. "Aspects of the Relationship Between Theories of Grammar and Theories of Processing." *Atlantis* 27 (1): 11–27.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge: The M.I.T. Press.

Friedmann, Na'ama. 2001. "Agrammatism and the Psychological Reality of the Syntactic Tree." *Journal of Psycholinguistic Research*. 30 (1): 71–90.

Harris, Randy. 1993. *The Linguistics Wars.* New York: Oxford University Press.

Hill, Archibald A., ed. 1962. *The Third Texas Conference on Problems of Linguistic Analysis in English: May 9–12, 1958*. Austin: University of Texas Press.

Katz, Jerrold J. and Jerry A. Fodor. 1963. "The Structure of a Semantic Theory." *Language* 39 (2): 170–210. New Jersey: Prentice-Hall.

Kennison, Shelia M. 2002. "Comprehending Noun Phrase Arguments and Adjuncts." *Journal of Psycholinguistic Research* 31 (1): 65–81.

van der Lely, Heather K. J., and Jackie Battell. 2003 "Wh-Movement in Children with Grammatical SLI: A Test of the RDDR Hypothesis." *Language* 79 (1): 153–181. New Jersey: Prentice-Hall.

van der Lely, Heather K. J. 1998. "SLI in Children: Movement, Economy, and Deficits in the Computational-Syntactic System." *Language Acquisition* 7 (4): 161–192.

Yang, Dan-Hee, Ik-Hwan Lee, and Pascual Cantos. 2002. "On the Corpus Size Needed for Compiling Comprehensive Computational Lexicon by Automatic Lexical Acquisistion" *Computers and the Humanities* 36 (2): 171–190.

Hedging in the Courtroom

This study examines the use and implications of hedging in the courtroom. The authors discuss the theory that those with the least power will hedge the most. To test this theory, the authors spent time observing the use of hedging by those in a courtroom in order to formulate informative conclusions.

Matthew LeGare & Candace Firl

Introduction

Pragmatics is a broad linguistic category that can be broken down into various aspects. One of these aspects is hedges. Hedges, which are used to mitigate the strength or power of a statement (Lewin 2001), have been studied quite thoroughly by linguists around the world through experiments in various venues and countries. Experiments dedicated to hedge use are found in areas such as newspaper discourse (Buitkiene 2008), dentistry discourse (Morales et al. 2008), scientific texts (Lewin 2001), gender (Tannen 1996), and culture (Vold 2006), just to name a few. Hedges have also been studied within the context of the courtroom. While this setting has been observed and tested with indirect relation to hedging, few studies have focused intensely on hedging alone and the frequency of use within power groups. This is where the focus of our project lies. First, however, a brief overview of work done with hedges in court.

Translation is one subject of study which has dealt with hedges being used in the courtroom. Susan Berk-Seligson completed a study which analyzed how statements change when translated for use in court (1987). She found that the statements in her experiment generally became longer after translation into the court language (English). She explained that this was because fragmented speech became a narrative. Her study hints at the main point of our experiment when it discusses the effects of powerlessness. Her experiment suggests that hedges are used more by witnesses and by others who are in a situation where they have little or no power. Interestingly, it may suggest that this is the case only in English since she entertains the idea that translators are the ones inserting hedges rather than translating them from the original language into English. Thus, without translation, it is possible that the witnesses do not use many hedges (or elements of powerlessness) and there really is no difference between them and the lawyers and judges. Furthermore, this experiment may even dispute the idea of hedges being used more frequently by witnesses and less frequently by judges in a courtroom

because it is only in the narrative that hedges are found and not in the courtroom discourse, despite power issues.

Research done on hedging in the courtroom generally focuses on the power variable, probably because the courtroom is a prime place to study power because of its being one of the only settings of absolute power. Few other places are set up so that all agree on power structure and allow for someone like a judge to rule without opposition. In a more direct study of this power, James J. Bradac, Michael R. Hemphill, and Charles H. Tardy (1981) studied the perception of defendants and plaintiffs in a courtroom situation. Powerless and powerful speech was created for the defendants and plaintiffs by inserting or removing such things as hedges, intensifiers, and hesitation forms. A third party then analyzed the defendants and plaintiffs according to their speech. Defendants and plaintiffs were both considered more blameworthy or violent when using powerful speech and more lenient when using powerless speech. This too relates to the study at hand because in this case, the image created was done by adding or subtracting features of speech. The image that a judge, a lawyer, or even a witness wants to have may affect his or her speech consciously, or it may make no difference (being solely related to the power structure) and be subconscious. Allen E. Lind and William M. O'Barr (1979) studied the use of pragmatic elements according to social power. In their transcriptions of trials they found that language features such as intensifiers, hedges, polite forms, hesitations forms, and deictic phrases were inversely related to social status and power. This study deals more directly with and acts as a precursor to the study at hand, where hedges will be studied specifically in relation to power. From Lind and O'Barr's study, it would be assumed that hedges will be found more frequently in the speech of those without power.

While research exists on hedges, though not specifically in relation to the power structure in the courtroom, little has been done looking specifically at hedges as a distinct element rather than as part of a group of powerless elements. Furthermore, studies have been

conducted in relation to perceived power of speakers and outsiders, but little, if any, research can be found comparing hedge frequency within the different power levels of court: that of judges, lawyers, and witnesses. Our project seeks to explore these venues.

Those with the Least Power Will Hedge the Most

Some studies, such as Hosman and Wright (1987), have found that speakers who speak in a powerful style are more attractive and credible than those who speak with a powerless style. The power scale in a courtroom is taken to be linear with a judge at the top, being the most powerful, followed by the lawyers, and finally the witnesses. For this reason, it would be hypothesized that judges and especially lawyers would try their hardest to delete all elements of powerless speech elements. One may think witnesses would also try to use the same powerful speech to convince the lawyers, judges, and, at times, juries. However, witnesses have much less time to practice and learn the art of powerful speech, and sometimes they do not want to sound certain because they are not certain, and they do not want to falsely accuse. With this being said, support leans toward fewer hedges being used by those with more power and more hedges being used by those with less power.

Other research has been done concerning qualifiers (another term for hedges and truthfulness). Some of this research concerns native and nonnative English speakers. Naugle (2010), for instance, found that lawyers who used qualifiers were seen as being more truthful and dependable. We similarly posit that this research indirectly shows that lawyers are more likely to use qualifiers in the courtroom than judges, if only to seem more personable to defendants and witnesses. This is because judges can come across however they please, but lawyers need to be liked by their clients for the sake of their livelihood.
Finally, it has been found that "hedges might undermine the argument

quality because hedges denote powerlessness, thereby weakening strong arguments" (Durik et al. 2008). Since lawyers must honor their subjection to the judge and witnesses are intrinsically uncomfortable and unfamiliar in a court setting, it is easy to deduce that both parties will use more hedges than the judge, who is subject to no one and entirely familiar with the setting. If this is the case, hedging could be definitively linked to power in the courtroom, and even outside of it, in a precise manner.

In essence, hedging makes speech seem powerless, which in turn affects the speaker's image. While the speaker may seem more personable, he or she may also come across as weaker. All of these characteristics have a place within the court system, where the power structure is linear. Within this structure, roles play out to support an increase in hedging frequency with less power and authority.

There Will Be No Considerable Difference

To contrast the aforementioned side, others would argue there is no fundamental difference in courtroom discourse. Although the courtroom is a different venue of speech with its own rules and dictates, it is unreasonable to believe that these rules would affect the intrinsic basis of power and power-negation in modern sociolinguistics. There have been venue-specific studies performed (Fahy 2010) in which no statistical significance was found.

Additionally, some would argue that hedges are not indicative of a lack of power. Dirk Hovey (2004) quotes Gudrun Clemen (1997) in expressing that "hedges need not necessarily have a weakening effect, but can as well enforce what is said." Thus, judges may be just as likely as a lawyer or witness to use hedges for emphasis. In this case, no correlation between power and hedging would be found because of equal use across the board. Furthermore, it has

been found that nonnative speakers of English do not differentiate between hedged versions of a statement and non-hedged versions of a statement by lawyers in court (Naugle 2010). Thus, people do not notice a difference in meaning when hedging is used. If this is the case, there would be no reason for any party to adapt their speech according to their role in a courtroom setting. These facts seem to support the notion that there is no appreciable difference or importance in hedge usage in the courtroom. This would imply that hedging is not tied to power and that it has no sociolinguistic bearing in the courtroom.

Methodology

For this experiment, about twelve hours total were spent observing court trials at the Provo City Courthouse in Provo, Utah. While listening to these trials, we recorded the use of hedges by the judge, lawyers, and witnesses. The hedges which were tallied and taken into account were pulled from Tim Rowland's and Ulugbek Nurmukhamedov and Soon Hyon Kim's experiments. The first hedges on the list we created (*I think*, *maybe*, *probably*, and *possibly*) stemmed from what Rowland (2000, 129) called plausibility shields. The rest of the hedges were taken from Nurmukhamedov and Kim's (2009, 274) modals of ability (*can, might, could be able to*), expressions of possibility (*it would be a good idea, it would be better, it would be nice, it is better, perhaps*), and personal attributions (*I believe, in my opinion*). These fourteen hedges constitute our list for this experiment.

Upon entering the courtroom, each character in the court was given a number and was kept track of by jotting down features. When a hedge was used, a tally mark was given to indicate which hedge was used, by which character, and to whom it was directed. The tally marks from all observation sessions were added to represent the total for each individual character and the group they belonged to (judges,

lawyers, and witnesses) and further broken down into groups addressing other groups (judges to lawyers, judges to witnesses, lawyers to judges, etc.). The amount of time each character spent speaking was recorded throughout each session in order to normalize the number of hedges used by each group per minute. Thus, the total number of hedges was divided by the time spoken and compared to the other groups' statistics. This allowed us to determine the difference in hedge use among the groups. For statistical analysis we conducted both an independent t-test of the two groups—colloquial and courtroom—and a multiple regression analysis to investigate the effect that hedges have in the courtroom.

For a control, we each watched an episode of *The Office* and recorded the hedges used by each person. This was done more accurately by using transcriptions found online. As we watched, we recorded the amount of time spoken by each character in order to replicate our experiment above. The power structure was taken into account in order to mimic the experiment in the courtroom, but only the colloquial level was used as a control. We used colloquial discourse for a control because colloquial speech is where most of the research on hedging has been performed and validated. For the purpose of this study we defined colloquial speech as speech between two people in the same level. For *The Office*, this translated into speech among nonmanagement employees.

The threshold of significance in this experiment is set at $p < 0.03$ (i.e., less than 3-percent chance due to random occurance). Furthermore, if our statistical analysis is greater than 0.03 and if the hedge per minute rate differs by 0.03 or more, our findings are significant toward the discourse and use of hedges according to power structure in the courtroom. Should hedges per minute and correlations be split in their significance, we will give precedence to the statistical analysis, as it can tell us quantitatively what is happening in the courtroom.

Analysis

Our analysis will be broken down into three sections. First we will discuss the control, followed by the courtroom, and finally the totals.

Control

From this individual breakdown in figure 1 we can compile the group statistics found below in figure 2:

Name	Minutes	Hedges	Hedges/Minute
Dwight	1.98	2	1.0084
Deryl	7.53	3	0.3982
Jim	0.98	5	5.1020
Mrs. California	2.85	6	2.1053
Oscar	0.25	0	0.0000
Phyllis	0.08	0	0.0000
Angela	0.10	0	0.0000
Kelly	0.05	0	0.0000
Toby	0.08	0	0.0000
Kevin	3.33	2	0.6001

Figure 1. Individual hedge usage in *The Office*, used as colloquial standard

Minutes	Hedges	Hedges/Minute
21:12	21	0.990566

Figure 2. Colloquial group hedge usage

We can see from figure 2 that the characters tend to use about one hedge per minute in colloquial speech. This was interpreted as the standard rate and used as the basis of hedge per minute comparisons, while the table in figure 1 was used to calculate independent t-test and multiple-regression analysis.

Courtroom

We will present the data from our courtroom observations from the top of the hierarchy to the bottom—that is, from judge to lawyer to witness.

	Minutes	**Hedges**	**Hedges/Minute**
Judge W to Witness	2.03	0	0.0000
Judge A to Witness	0.25	0	0.0000

Figure 3. Judge-to-witness discourse

	Minutes	**Hedges**	**Hedges/Minute**
Witness W to Judge	0.78	0	0.0000
Witness T to Judge	0.5	0	0.0000

Figure 4. Witness-to-judge discourse

As shown in the charts above, we were unable to collect any instances of a judge using a hedge when speaking to a witness or vice versa. This is an unfortunate anomaly in our study, one which will be discussed more thoroughly in the conclusion.

	Minutes	**Hedges**	**Hedges/Minute**
Judge Z to Lawyer	3.00	0	0.0000
Judge Y to Lawyer	6.75	0	0.0000
Judge X to Lawyer	3.50	1	0.2857
Judge A to Lawyer	0.50	0	0.0000
Judge A to Lawyer (1)	2.08	1	0.4800
Judge A to Lawyer	0.08	0	0.0000

Figure 5. Judge-to-lawyer discourse

We can see here that judges rarely use hedges when talking to lawyers. This was to be expected; Hosman and Wright (1987) referenced that the only times hedges were used were likely as power mitigators when talking to lawyers.

	Minutes	Hedges	Hedges/Minute
Lawyer Z to Judge	1.20	2	1.6667
Lawyer X to Judge	1.33	1	0.7500
Lawyer W to Judge	2.00	0	0.7500
Lawyer V to Judge	0.25	0	0.0000
Lawyer U to Judge	3.00	2	0.0000
Lawyer T to Judge	1.00	1	0.6667
Lawyer S to Judge	0.08	0	1.0000
Lawyer R to Judge	0.75	1	0.0000
Lawyer Q to Judge	4.25	3	1.3333
Lawyer P to Judge	3.67	1	0.7059
Lawyer O to Judge	0.17	0	0.2727
Lawyer M to Judge	3.50	1	0.2857
Lawyer M to Judge (1)	0.333	0	0.0000
Lawyer M to Judge (2)	1.0833	3	2.7693
Lawyer L to Judge	0.50	1	2.0000

Figure 6. Lawyer-to-judge discourse

Here we can see that lawyers, on the whole, tend to use more hedges when talking to judges than vice versa. Additionally, we can see an increase in rows of data in comparison to the previous charts. This is due to the fact that lawyers must consult with the judge on many procedural issues, such as approaching the witness, presenting an exhibit, and getting an unmarked exhibit marked. One must also consider the fact that there are always more lawyers than judges in a courtroom, which further compounds the previously stated facts.

	Minutes	Hedges	Hedges/Minute
Lawyer Z to Witness	11.97	1	0.0836
Lawyer Y to Witness	25.00	2	0.0800
Lawyer M to Witness (0)	13.50	0	0.0000
Lawyer M to Witness (1)	4.02	1	0.2488
Lawyer L to Witness	0.58	0	0.0000

	(Minutes)	(Hedges)	(Hedges/Minute)
Lawyer L to Witness (1)	1.83	1	0.5455
Lawyer L to Witness (2)	0.17	0	0.0000
Lawyer L to Witness (3)	1.67	0	0.0000
Lawyer L to Witness	1.75	0	0.0000

Figure 7. Lawyer-to-witness discourse

Much as when judges talk to lawyers, we can see here that lawyers tend to use few hedges when speaking with witnesses, as hypothesized.

	Minutes	Hedges	Hedges/Minute
Witness Z to Lawyer	43.00	34	0.7907
Witness Y to Lawyer	4.50	10	2.2222
Witness X to Lawyer	3.20	7	2.1875
Witness U to Lawyer	6.50	4	0.6154
Witness T to Lawyer	17.58	12	0.6825
Witness S to Lawyer	7.58	7	0.9231
Witness S to Lawyer (1)	1.33	5	3.7594
Witness R to Lawyer	4.50	9	2.0000
Witness Q to Lawyer	6.25	10	1.6000

Figure 8. Witness-to-lawyer discourse

Finally, witnesses tend to hedge at or above the standard rate of 0.99 hedges per minute. This could be due to a variety of factors such as anxiety, courtroom familiarity, or any type of courtroom training received. But on the whole the witnesses tend to perform at about the standard rate.

From the total charts we can quickly see that the witness group, when speaking to lawyers, comes closest to matching the colloquial rate of hedges per minute. The judge, as postulated, rarely uses hedges in the courtroom. From our methodology we see that there are significant differences in hedge-per-minute rates for lawyers and judges in the courtroom; both groups fall far below the standard 0.99 hedges per minute.

Courtroom	Total Time	Total Hedges	Hedges/Minute
L to W	2.65	5	0.08
L to J	0.65	4	0.26
W to L	3.75	98	1.09
W to J	0.12	0	0
J to W	0.19	0	0
J to L	0.68	2	0.12

Figure 9. Courtroom group totals

Time	Hedges	Hedges/Minute
21:12	21	0.99

Figure 10. Colloquial group totals

When we ran our independent t-test and multiple regression analysis, we found that there is almost literally no correlation between the variables we observed, due to the insufficient size of our discourse sample. From the charts above we can see that the general trend is for witnesses to hedge at a rate higher than judges and lawyers in the courtroom; however, more research and much more data will need to be compiled before any concrete, statistically significant claims can be made.

Despite the fact that our hedges-per-minute measurement met our stated criterion for significance, our results support our null hypothesis, that there is no measurable difference in the courtroom between who hedges the most and who hedges the least, when compared with the colloquial rate. This is because our statistical analysis yielded no statistically significant results, which would have been much more definitive statements than our qualitative hedge-per-minute ratio. While this result was surprising to the researchers, we acknowledge that a much more robust study will be needed to fully investigate the matter.

Conclusion

This study has concluded that while some sort of trend is occurring in the courtroom with respect to hedges, none of our findings were statistically significant, most likely due to a simple lack of data. What we can see is that judges and lawyers definitely hedge at a rate far removed from the colloquial average. However, because of the lack of concrete correlations, we were not able to link hedge rates in and out of the courtroom. Apart from our lack of data, this could be due to a variety of variables, from demographic and socioeconomic variables to the fact that two completely different registers are being used in and out of the courtroom.

In the scope of sociolinguistics, the observed trend that judges and lawyers hedge at a rate far less frequent than witnesses is an important, if not unexpected, discovery. Previous forensic studies which have taken the absolute power structure of the courtroom for granted can relax knowing that their assumptions have been validated in the courtroom to some extent, although not in a statistically significant way.

The researchers must acknowledge openly both the plethora of variables and constraints which were not in our power to control and the fact that a much larger data set will be necessary to validate the patterns found in this study. Working in the realm of the courtroom, we had no power over age, sex, socioeconomic status, education, or any other demographic type variables. While we are grateful for the opportunity to observe courtroom cases, future studies may wish to try to control these variables. Additionally, we could not control the volume of speech produced by any one individual. Some cases would see lawyers doing the majority of the talking, others the witnesses, and yet others had defendants defending themselves. This is most evident in our limited pool of judges' speech to analyze and in the fact that not a single judge or witness used a hedge when addressing the other.

Due to time and personel constraints, the researchers were not able to conduct a proper study of colloquial speech and had to settle

for a TV approximation via episodes of the popular television show *The Office.* Had time and means been available, a full-fledged study of hedges both in and out of the courtroom would have been preferred.

Despite these obvious flaws in the research, we still believe it to be a pertinent and necessary attempt to further our knowledge of how hedges are used in the courtroom. This research could be used to have a more in-depth study of power and solidarity, as there is no other place in the world with such a clear-cut power hierarchy. The research could also be used to validate and further test current theories on hedge usage.

Future Work

For future work we feel that a number of things could be improved. First, we would like to see many of the uncontrolled variables controlled. While such control could not be asserted in a standard courtroom, alternative methods could be used to approximate the proceedings of an actual trial. Various mock trial clubs at high schools and colleges across the country—or perhaps actual law students—could be used in this endeavor. In a mock trial setup, researchers could more easily control demographic variables such as age, gender, and race. By controlling variables in a reliable way, the results of this study could be further validated and improved. Additionally, a mock trial setting would allow researchers to have slightly more control of turn-taking and dictating how much certain key players talk in the courtroom. As competitive mock trial cases are often presided over by an actual judge, these cases could provide, if nothing else, a more robust amount of data concerning judges in the court setting.

The researchers are also in favor of a more robust study into the matter of hedges in the courtroom. From the results we can see that a far greater number of hours than we were able to observe will be necessary to further understand the courtroom phenomena at a statistically significant level. Additionally, due to time and personel

constraints, a full-fledged paired group study of both the colloquial and the courtroom discourse with relation to hedges was not possible. Since courtroom and colloquial discourse are, in general, fluid acts with many variables, many more studies should be conducted on the matter with an increase in both scope and magnitude.

One future study could focus on correlating or comparing courtroom and actual colloquial discourse instead of pre-scripted, popular media discourse. Naturally, future studies—especially something to this degree—would require a greater number of hours in both settings, done preferably by a whole team of researchers. Another future study would include tracking hedge use between genders. It would be interesting to see if one gender hedges more than the other in the courtroom. This would provide further strength to the results because it would throw gender out whereas in the study at hand, a majority of lawyers were male and the witnesses were balanced between male and female.

While we were observing in court, a couple of ideas concerning the witnesses came to mind for future work as well. The first idea deals with socioeconomic status and familiarity with the legal system. While some witnesses were clearly more comfortable and also unwilling to subject themselves to the power built into the judicial system, others seemed incredibly out of place and at the bottom of the power ladder, even outside of court. Thus, it would be interesting to see if hedging was statistically different, or even significant, across different types of witnesses.

Second, witnesses are often prepared to answer the questions asked them by their own lawyers. In essence, they will prepare scripts to answer these questions. They do not, however, have the opportunity to practice what they will say in response to the opposing lawyers' questions. With this situation, an experiment could be done concerning hedging frequency between scripted and non-scripted speech. This, of course, would have more complications than discussed here because of the difficulty of defining "scripted" and "non-scripted" speech as well

as knowing what has been practiced versus what hasn't been; such a study would require a great amount of support from the court. Nevertheless, it presents an intriguing project idea.

In addition to a more robust study, we feel that this same methodology could be applied to political arenas, such as the Senate and State Senate, to great effect. Preliminary work has been done in this field in Spanish by P. E. Acosta (2009), but no such studies have been conducted concerning the US Senate or English Parliament. Being admittedly biased toward this particular topic, we feel that future work in the areas suggested could help further the field of sociolinguistics as a whole with particular regard to hedges. This type of hedge research would, in turn, shed more light on the core issue of power vs. solidarity, which is a topic of great interest in our field today.

References

Acosta, Pablo Enrique. "Gender Differences in the Use of Hedges as a Politeness Strategy During the Debates of the Colombian Senate." MA thesis, Ball State University. 2009. <http://cardinalscholar.bsu.edu/bitstream/123456789/193610/1/Pacostaacosta_2009-1_BODY.pdf>

Berk-Seligson, Susan. 1987. "The Intersection of Testimony Styles in Interpreted Judicial Proceedings: Pragmatic Alterations in Spanish Testimony." *Linguistics* 25:1087–1125

Bradac, James J., Michael R. Hemphill, and Charles H. Tardy. 1981. "Language Style on Trial: Effects of 'Powerful' and 'Powerless' Speech Upon Judgments of Victims and Villains." *Western Journal of Speech Communication* 45:327–341.

Buitkiene, Janina. 2008. "Hedging in Newspaper Discourse." *Zmogus irZodis* 10:11–15.

Durik, Amanda M., M. Anne Britt, Rebecca Reynolds, and Jennifer Storey. 2008. "The Effects of Hedges in Persuasive Arguments: a Nuanced Analysis of Language." *Journal of Language and Social Psychology* 27:217–234.

Fahy, Patrick. 2002. "Use of Linguistic Qualifiers and Intensifiers in a Computer Conference." *American Journal of Distance Education* 1:5–22.

Hosman, L. A., and J. W. Wright, II. 1987. "The Effects of Hedges and Hesitations on Impression Formation in a Simulated Courtroom Context." *Western Journal of Speech Communication* 51:173–188.

Hovy, Dirk. 2004. "Hedges." <http://www.dirkhovy.com/portfolio/papers/download/hedges_dirk_hovy.pdf>

Lewin, Beverly. 2001. "From hedging to heightening: toning down and up in scientific texts." *Malbourne Papers in Linguistics and Applied Linguistics* 1:17–28.

Lind, E. Allen and William M. O'Barr. 1979. "Social Significance of Speech in the Courtroom". *In Language and Social Psychology,* 66–87. Baltimore: University Park Press.

Morales, Oscar Alberto, Daniel Cassany, Ernesto Ilich Marin-Altuve, and Carolina Gonzalez-Pena. 2008. "Hedging in the Discourse of Dentistry: Clinical Cases in Spanish Journals." *Circulo de linguistica aplicada a la comunicacion* 34:18–41.

Naugle, William. "Native and Nonnative English Speakers' Perception of Hedging in the Oral Argument of Civil Rights Cases." PhD diss., New York University. 2011. <http://gradworks.umi.com/3454485.pdf>

Nurmukhamedov, Ulugbek and Soon Hyon Kim. 2009. "'Would you perhaps consider . . . ': Hedged Comments in ESL Writing." *ELT Journal* 64:272–282.

Rowland, Tim. 2000. *The Pragmatics of Mathematics Education: Vagueness in Mathematical Discourse.* London: Falmer Press.

Tannen, Deborah. 1996. "The Sex-class Linked Framing of Talk at Work." *In Gender and Discourse* 195–218. New York: Oxford UP.

Vold, Eva. 2006. "Epistemic Modality Markers in Research Articles: A Cross-linguistic and Cross-disciplinary Study." *International Journal of Applied Linguistics* 16:61–87.

Peircean Thirdness and Text Linguistics: The Effect of Sense-Lines on Coherence and Comprehension

The author examines Charles S. Peirce's concept of "Thirdness" and its significance in reading comprehension. More specifically, the goal of this study was to determine whether or not the use of sense-lines (an implementation of Thirdness) increases understanding. To test this, the author extracted reading prompts from the Graduate Record Examination (GRE), reformatted them with sense-lines, and administered comprehension quizzes.

Stetson Robinson

Introduction

Philosopher, mathematician, and scientist Charles S. Peirce (1839–1914) professed that our understanding of the world around us is rooted in three different "categories" of relational existence: Firstness, Secondness, and Thirdness (Peirce 1868). According to Peirce, Firstness is the existence of a subject "without regard to anything else"; Secondness is the existence of a subject "with respect to a second but without any third"; and Thirdness is the existence of a subject "in bringing a second and third into relation to each other" (Dictionary of Peirce's Terms s.v. "Firstness").

Although these concepts are abstract by nature, they provide enthralling insight when considered in a linguistic paradigm, specifically in text linguistics. For instance, consider the work of a text editor as she pores over the text and corrects a few misspelled words or adjusts the size and font of a footnote. She then recasts a compound sentence for parallelism to aid the reader in understanding the verb-noun relationship. And finally she molds the text into a template with a desired paragraphing style. In summary, she has adjusted the manuscript to conform to all three Peircean categories. To be clear, we can use this anecdote to define these categories of existence as they apply specifically to textual design: "Firstness requires both unity and variety" (fonts, colors, orthography, etc.) for intelligibility; "Secondness requires contrast and action" (cues, links, syntactic parallelism, contextual appropriateness, etc.) for cohesion; and "Thirdness requires pattern" (hierarchy, punctuation, layout, etc.) to facilitate overall coherence (Manning 2011). Although all three categories of existence are each compelling areas of study, this investigation focuses specifically on Peircean Thirdness as it relates to sense-lines and its influence on overall coherence of alterable prose (i.e., prose that can be or already is adjusted for Firstness and Secondness).

Before the conventions of punctuation were adopted, a more primitive way of dividing text into more readable units was to break the

text into separate lines of thoughts or ideas (usually clauses or phrases). This method was essentially the only way to edit a manuscript, primarily because most ideas worthy of being written were otherwise unalterable (e.g., recordings of the dictations either of spiritual leaders declaring the word of God or of secular leaders declaring the law of the land). Ancient monks in the scriptorium, for instance, used this very method of writing (which they referred to as *per cola et commata*) to more clearly record dictated utterances (Manguel 1996). As a result of the unalterable nature of recorded speech, ancient record-keepers did not have much freedom to adjust the text for grammar or mechanics (i.e., Firstness or Secondness), thus Thirdness played an essential role in writing and editing.

Today editors and writers likewise employ principles of Thirdness with text that cannot be altered for Firstness or Secondness because the author is dead or otherwise unable (or unwilling) to authorize alteration. A prominent example of this is scripture. Bradbury Thompson's *Washburn College Bible* employs the use of sense-lines "for easier reading and better comprehension" (Thompson 1979). Additionally, writer and linguist Royal Skousen of Brigham Young University created a sense-line version of the Book of Mormon, an ancient book of scripture from the American continent, to assist readers in understanding an older style of English. In his text, Skousen offers his own explanation as to why the sense-line format is so helpful for readability in ancient text:

> Sense-lines can assist readers in differentiating phrases and clauses, identifying constituent grammatical units, and keeping track of subjects, main verbs, and modifiers. For instance, sense-lines can be helpful in interpreting cases where a prepositional phrase is displaced from its expected position, as in Mosiah 26:23:
>
> > And it is I that granteth unto him that believeth
> > in the end a place at my right hand.
>
> Here the phrase "in the end" refers to the verb *granteth* rather than *believeth*. . . . Reordering the sentence makes its meaning clearer than punctuation could. (*The Book of Mormon: The Earliest Text,* xliii)

The primary disadvantages of sense-lines and other areas of editing that adjust for Thirdness are paper-space requirements and impracticality. It is more practical and saves more space on a page to run text together in conventional paragraphing—sense-lines in particular require more space and therefore more paper (and thus more money) than standard paragraphing. Therefore, for practicality and space, among other reasons, Thirdness plays an inferior role than it otherwise would in alterable text, which editors are allowed and supposed to adjust for Firstness and Secondness. (Note: For simplicity and brevity, from this point on in the context of this article the verb *edit* will refer to altering a text for Firstness and Secondness, and the term *sense-lines* will represent the principle of Thirdness as a whole as it compares to the effects of Firstness and Secondness.)

The question, now, is whether Thirdness, specifically as it relates to sense-lines, will actually improve cohesion to the point of increased readability and therefore comprehension of text that has already been altered for grammar, mechanics, and other areas of Firstness and Secondness. In other words, will a sense-line format improve reading comprehension of text already edited for clarity? Or will said format have no effect on reading comprehension because the existing alterations make sense-lines obsolete (i.e., Firstness and Secondness overrule Thirdness)?

Sense-Line Format Will Not Affect Comprehension

This side of the issue is the null hypothesis; that is, using sense-lines in edited prose will have no effect on readability. A study by Chiu (2004) found that a Taiwanese student learning English to get a master's degree in English literature was aided significantly by syntactic considerations such as global English (Secondness), but no improvement was seen as a result of different styles of paragraphing and layout that were introduced to the student. Others like Conner (1996) and Schneider (1991)

have found similar trends with both ESL and native English speakers. This suggests that an accommodation like sense-lines would be relatively ineffective on the readability of prose that has already been edited for semantic appropriateness and syntactic uniformity.

Sense-Line Format Will Improve Comprehension

This side of the issue holds that accommodating text with sense-lines will indeed increase readability and comprehension because it aids the reader along a topical progression that is natural for and familiar to the human mind. Kopple (1989) would support such a claim since he believed and outlined that part of the functional role of text is to cater to immediate visual perception of the prose. Odler (1989) and Manguel (1996) would also agree with this claim, but more specifically that text format increases reading comprehension for both ESL learners and native English speakers alike, implying that text format has a significant role in second-language acquisition (a topic meriting volumes of discussion, but unfortunately not in this investigation). This side suggests, then, that principles of Thirdness relating to sense-lines are tremendously underused and would significantly improve readability and comprehension of edited texts.

Methodology

The basic objective of the methodology is to compare the results of randomly selected participants who took two quizzes with identical questions and with an identical reading prompt, the only differing variable being that one reading prompt was formatted with sense-lines and the other was formatted with conventional paragraphing (see appendix).

Stimuli

The reading prompt and comprehension questions used for the quiz were taken directly from the verbal portion of a previous Graduate Record Examination (GRE) available online (JumboTests 2011). Reading prompts from the GRE are near-perfect examples of edited text because they are selected and edited extensively by highly trained professionals in correspondence with standardized testing requirements—that is, exam excerpts are required to be clear and unambiguous while still challenging enough in content to target a specific level of intellect. Likewise, the comprehension questions themselves on the GRE receive the same intense treatment for quality and fairness in assessing understanding of the given prompt.

The specific reading prompt selected for this study was approximately one hundred words in length, which is long enough to provide sufficient data to be tested but short enough to not fatigue the reader. Additionally, this prompt was chosen because the content did not require specialization in any specific area of study in order to understand the material.

Five questions were used to test the reader's comprehension of the provided reading prompt. Four of the five comprehension questions appeared precisely as they did on the actual GRE; therefore, the fairness and validity of the questions is not an influential variable as far as this investigation is concerned. The fifth question was a qualitative question that asked the readers how well they *thought* they performed on the quiz; this information was used to measure predictability of success as well as patterns in self-confidence (which will be discussed below).

And finally, the most important aspect of the stimuli, as mentioned above, is the variation in layout between the two. The reading prompt of one quiz was formatted in conventional paragraph form while the other was formatted with sense-lines. The quizzes were administered via an online survey that was sent to participants electronically. There was no time limit, but the quizzes were designed such that the readers

could not return to the reading prompt as a reference once they moved on to the questions, and they likewise could not return to previously answered questions. Brief instructions were also included at the beginning of the quiz explaining the above conditions as well as the readers' objective to perform to the best of their ability.

Subjects

Two groups of fifty participants (total of one hundred) completed the online quizzes. One group took the quiz with the conventionally paragraphed reading prompt while the other took the quiz with the prompt formatted with sense-lines. Each of the two groups was selected such that gender and age were consistent in both—specifically, each group consisted of twenty-five females and twenty-five males, and all participants were between ages eighteen and thirty (ages within the intellectual umbrella of the GRE). All participants reported that they had no unique circumstances of mental disability.

Additionally, the participants were unaware of the true purpose of the quiz (i.e., they did not know that this was an exercise to test the format of text), but were instructed simply to take the quiz and perform as well as they could.

Calculations

Once participants completed their respective quizzes, the responses were collected to create two data sets reflecting the performance of each group. One data set represented the percentage of how many participants in one group answered each individual question correctly, and the other data set represented the percentage of how many participants performed worse than, the same as, or better than they predicated (recall the fifth question of the quiz that asked how many of the four questions they *thought* they answered correctly). With these data sets in place for each of the two groups, it was possible to perform a comparative analysis and draw conclusions from the data.

Analysis

Below (figure 1) are the results from the quizzes. The comparative calculations are divided into two sections of respective areas of measurement. The first section compares both groups in terms of the percentage of group members who answered a given question correctly. The second section compares both groups in terms of the percentage of group members that performed worse than, the same as, or better than they predicted.

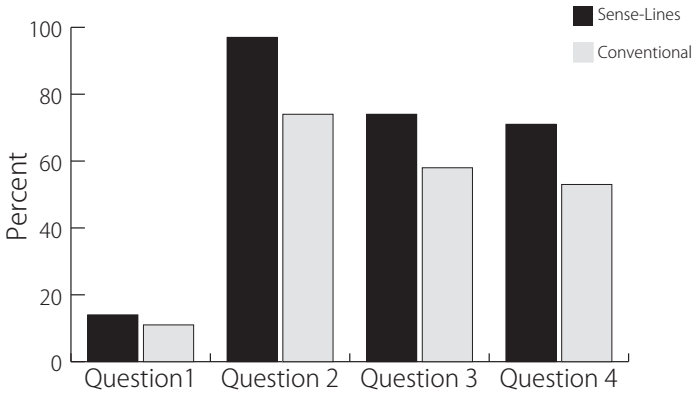


Figure 1. Percentage of participants per group who answered each question correctly

Comparative Accuracy

The group whose reading prompt was formatted conventionally performed worse on every question than the group whose prompt was formatted in sense-lines. Of the conventional group, 11 percent answered question 1 correctly compared to 14 percent from the sense-line group. On question 2, the majority of both groups' members answered correctly, but the conventional group only had 74 percent who selected correctly while the sense-line group had a remarkable 97 percent. On average, the sense-line group performed 14 percent better on every question than the conventional group.

The data above also show that each group followed a similar trend of the number of correct answers per question, only the sense-line group was on a consistently higher plane. For example, from question 2 to question 3 there was approximately the same percent decrease in correct answers for both groups, which was about a 20-percent decrease. In other words, 20 percent fewer readers in both groups answered question 3 correctly than answered question 2. The same trend appears from question 1 to 2 and from question 3 to 4. This pattern is interesting because it shows that the difficulty level from question to question was proportionate for both groups. Therefore, we know that the validity and fairness of the questions is not influential on the results and conclusion that the sense-line group performed better on the quiz as a whole. These findings support the claim that sense-line formatting does in fact improve comprehension. (Also, an interesting outcome not reflected in figure 1 is that three times more group members from the sense-line group answered all of the questions correctly than the non–sense-line group.)

Additionally, I performed a t-test to find a concrete threshold of significance and to be positive that the comparative outcome of the two groups was not simply a result of some random series of influential events. The results of my t-test yielded the following data:

$$t(98) = -6.962;\ p < 0.0005$$

From these results I calculated that the percent-chance that the comparative outcome of the two groups was random was less than 0.05 percent. From the same t-test, I found that the sense-line group had a 68-percent rate of accuracy while the conventional group had only a 35-percent accuracy rate. In summary, the chance that these data represent a random outcome is miniscule and the 33-percent difference is significant.

Comparative Confidence

The fifth question of the quiz asked participants how many of the four questions they predicted that they answered correctly. Of the participants in the conventional group, 73 percent performed worse on the

quiz than they thought (i.e., they answered fewer answers correctly than they predicted), 18 percent performed the same as they predicted, and 9 percent performed better than they predicted.

On the other hand, of the participants in the sense-line group, only 29 percent performed worse than they predicted, 52 percent performed the same as they predicted, and 19 percent performed better than predicted. (See figure 2.)

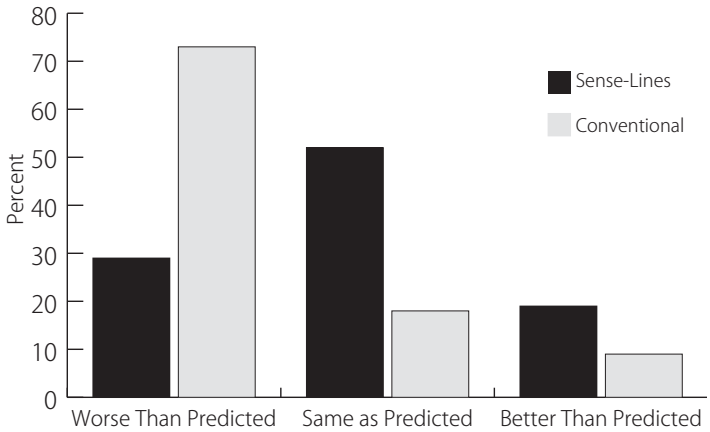


Figure 2. Participants' performance with respect to predicted performance

From this comparison, we see that the majority of the readers exposed to the conventional text format were overconfident in their comprehension of the text. The majority of the sense-line group, however, was neither overconfident nor lacking confidence, but rather performed exactly as they had predicted, which was significantly better on every question than those in the conventional group.

The fact that the majority of the sense-line group precisely predicted their results is intriguing. If the majority of the sense-line group had been overconfident, then the comparison between the two groups would have been obsolete in the context of this investigation. On the other hand, had the majority of the sense-line group lacked confidence

(i.e., performed better than they predicted), then this investigation would show that sense-lines improve comprehension but primarily on a subconscious level. However, because these data show that the majority of these group members performed precisely as they predicted, we know that sense-line format improves readability in a cognitive way such that readers not only understand the text better than they would otherwise, but they are also consciously aware of their understanding. This is very convincing evidence in support of the claim that sense-lines, and Thirdness in general, improve comprehension.

Conclusions

The data and discussion above lead us to conclude that the hypothesis that sense-lines significantly improve reading comprehension of previously edited text is accurate. We can likewise conclude that Thirdness still plays an essential role in the readability of text, even after the text has been conformed to Firstness and Secondness.

Furthermore, this study brings up important implications that need be mentioned. For instance, if in fact there is a correlation between sense-lines (or Thirdness in general), confidence level, and overall comprehension of information (as we observed from the data), then standardized test scores in general could have a gaping hole in validity (i.e., scores do not fairly and accurately represent the skills or knowledge tested) if a reading comprehension portion is required. The format of the reading comprehension portion of such exams could be structured such that test-takers are unable to reach their full potential of comprehension and confidence. If this is the case, then undergraduate or graduate schools that heavily base acceptance of students on standardized test scores that include reading portions (e.g., SAT, GRE, GMAT, LSAT, MCAT, etc.) are following a highly flawed system that unfairly sifts through applicants. Standardized-test writers and editors need not, of course, adopt specifically a sense-line

format on tests. Rather, the use of sense-lines in this investigation represents the potential defects of standard paragraphing and serves to raise questions about the current formatting of standardized tests, in an effort to increase awareness and perhaps propose more effective formatting alternatives.

There are, of course, limitations on this particular investigation given the scope of the research. First, the number of participants was small. Our pool consisted of two groups composed of 50 participants each with an equal gender division and a specific age range. This pool is decent for the scope of this investigation, but a much larger pool of individuals (at least one thousand) would yield a much more persuasive outcome. Also, only having one pool means only having one quiz (which is duplicated among the two groups). Having multiple pools (each still divided in two comparative groups, of course), however, allows for multiple quizzes and thus multiple genres of reading prompts. This verifies that the trends found are independent of the content of the text itself or the quiz questions.

An additional limitation on this study is that only sense-line formatting was used to represent the effects of Thirdness in general, which is admittedly a lofty representation. A study similar to this one would need to be conducted using an alternative to paragraphing other than sense-lines. This would confirm the claim that Thirdness in general, rather than sense-lines in particular, is highly influential on textual coherence and processing.

Future Work

With several limitations of this investigation outlined above, we can pinpoint specific areas of this study that merit further research. First of all, much more careful attention needs to be given to the formatting and types of questions that result in varying scores. For example, this study used four questions to test reader comprehension with no specific

accommodations for the delivery or format of those questions. If the format of the reading prompt had such a drastic influence on the readers, then there could very well be a similar trend if similar accommodations were made with the actual questions. Perhaps instead of having a vertical list of multiple choice answers that readers choose from, a horizontal list of options could be provided to test respondent accuracy. Conclusions drawn from a study like this would significantly improve the logic of test creation.

On the same note, because there were only four questions on the quiz, there are no data to support that readers would perform equally well on a much longer but similarly formatted exam as they did on this shorter quiz. In other words, could the duration of the exam make formatting adjustments obsolete? And regarding duration, would we find the same results if the reading prompt were much longer than the one provided for this investigation?

A slightly different methodology would be needed in order to accommodate the adjustments mentioned above. For example, there would have to be a different reading prompt used for the quiz, one that is much longer, perhaps five to six hundred words. Additionally, more questions would need to be added to test for fatigue and decay in reader response. If we wanted to test for formatting of the questions themselves, duplicates of each respectively formatted quiz would need to be made with contrasting formats of the actual questions.

An additional area of future work would be to run a similar test focused on gender and/or age influences on the results. A pool of participants similar to this study's could be drafted, but instead of testing groups with equal numbers of each gender, groups could be composed entirely of females or males, or entirely of people in one age bracket. A study with pools like these could yield very useful information about the role of gender and/or age in relation to Thirdness. Specifically, we could approach more psychological issues dealing with the reading and logical processing of female and male brains, or of certain age groups, giving us a helpful direction for educational

improvement and other academic considerations. It would additionally give us a better idea of how to cater to a global audience in our writing, which would be especially important in technical writing fields such as instruction manuals, language materials, or other areas where reader comprehension is critical.

References

Chiu, Yi-hui Florence. 2004. "Coaching a Student to Develop Coherence Based upon Topical Structure Analysis: A Case Study." *Journal of Language and Learning* 2 (2):154–170.

Connor, Ulla. 1996. *Contrastive Rhetoric: Cross-cultural Aspects of Second Language Writing.* Cambridge: Cambridge University Press.

JumboTests. 2011. "GRE Reading Comprehension Practice." <http://gre.jumbotests.com/tests/gre-reading-comprehension-practice>

Kopple, William J. Vande. 1989. *Clear and Coherent Prose: A Functional Approach*. Glenview: Scott Foresman and Co.

Manguel, Alberto. 1996. *A History of Reading*. Middlesex: Penguin.

Manning, Alan. 2011. "Assessing Visual Design: An Overview." <http://linguistics.byu.edu/classes/elang410am/overview.html>

Odlin, Terence. 1989. *Language Transfer: Cross-Linguistic Influence in Language Learning.* Cambridge: Cambridge University Press.

Peirce, Charles. 1868. "On a New List of Categories." *Proceedings of the American Academy of the Arts and Sciences* 7: 287–298.

Schneider, M., and U. Connor. 1991. "Analyzing Topical Structure in ESL Essays: Not All Topics are Equal." *Studies in Second Language Acquisition* 12: 411–427.

Skousen, Royal. 2009. *The Book of Mormon: The Earliest Text.* New Haven: Yale University Press.

Thompson, Bradbury. 1979. *The Washburn College Bible.* New York: Oxford University Press.

Appendix

Reading prompt one (conventional paragraphing)

> But man is not destined to vanish. He can be killed, but he cannot be destroyed, because his soul is deathless and his spirit is irrepressible. Therefore, though the situation seems dark in the context of the confrontation between the superpowers, the silver lining is provided by amazing phenomenon that the very nations which have spent incalculable resources and energy for the production of deadly weapons are desperately trying to find out how they might never be used. They threaten each other, intimidate each other, and go to the brink, but before the total hour arrives they withdraw from the brink.

Reading prompt two (sense-line formatting)

> But man is not destined to vanish.
> He can be killed,
> but he cannot be destroyed,
> because his soul is deathless and his spirit is irrepressible.
> Therefore,
> though the situation seems dark in the context of the confrontation between the superpowers,
> the silver lining is provided by amazing phenomenon
> that the very nations
> which have spent incalculable resources and energy for the production of deadly weapons
> are desperately trying to find out how they might never be used.
> They threaten each other, intimidate each other, and go to the brink, but before the total hour arrives
> they withdraw from the brink.

Multiple-choice questions that accompanied each separate prompt—bolded choice is the correct answer, according to GRE scoring guide

1. The main point from the author's view is that
   A) Man's soul and spirit are immortal.
   B) Man's destiny is not fully clear or visible.
   C) Human society will survive despite the serious threat of total annihilation.
   D) Man's safety is assured by the delicate balance of power in terms of nuclear weapons.
   E) Man's soul and spirit cannot be destroyed by superpowers.
2. The phrase "go to the brink" in the passage means
   A) Retreating from extreme danger.
   B) Declare war on each other.
   C) Advancing to the stage of war but not engaging in it.
   D) Negotiate for peace.
   E) Commit suicide.
3. In the author's opinion
   A) Mankind is heading towards complete destruction.
   B) There is a silver lining over the production of deadly weapons.
   C) Nations in possession of huge stockpiles of lethal weapons are trying hard to avoid actual conflict.
   D) Superpowers have at last realized the need for abandoning the production of lethal weapons.
   E) Huge stockpiles of destructive weapons have so far saved mankind from a catastrophe.
4. A suitable title for the above passage is
   A) Mounting Cost of Modern Weapons
   B) Destruction of Mankind Is Inevitable
   C) Man's Desire to Survive Inhibits Use of Deadly Weapons

D) Threats and Intimidation between Superpowers

E) Cowardly Retreat by Man

5. How many of the above questions do you think you answered correctly?

A) 1 question

B) 2 questions

C) 3 questions

D) All questions

How Aibileen Be: A Linguistic Analysis of Kathryn Stockett's Use of Black Vernacular English in *The Help*

The Help, *a best-selling novel featuring several African American protagonists, has been criticized for its inaccurate use of Black Vernacular English (BVE). This article is an analysis of how BVE is used in the novel, with the end of determining whether or not these criticisms are correct.*

Michael Wyatt

Introduction

Kathryn Stockett's 2009 novel, *The Help*, has been the subject of much sociological, literary, and linguistic controversy. With the recent release of the 2011 film adaptation, increasing numbers of scholars and critics have added their voices to the heated discourse concerning the novel's ethics and accuracy. While resonating deeply with both white and black people, the book and film have simultaneously infuriated individuals of both ethnicities.

One of the key issues surrounding the novel is its use of Black Vernacular English (BVE).[1] The book is told from the perspectives of three women in a stream-of-consciousness narrative. Two of these narrators are black domestic workers from Jackson, Mississippi whose voices are written in BVE. Kathryn Stockett herself is a white woman from Jackson, Mississippi. In an essay at the end of *The Help*, "Too Little, Too Late," Stockett explains that a black domestic helped raise her during her childhood, allowing her ample time to hear BVE during her youth. Also, Stockett's brother's maid, Ablene Cooper, is currently seeking legal action against Stockett for using her likeness for the fictional Aibileen (Blakely 2011). Perhaps Stockett is also using her linguistic likeness.

The accuracy of Stockett's depiction of BVE has been both lauded and ridiculed. Although critics and scholars from a host of disciplines have found fertile ground for contention and debate in *The Help*, the book has yet to be examined by a linguist. In an effort to examine the accuracy of BVE in *The Help*, I will primarily study Aibileen's passages. Of all the characters, she has the most marked dialect, so if the accusation of caricature of BVE is valid, it will find its indictment in her voice.

"Pitch-Perfect Voices"

The dust jacket of *The Help* claims, "In pitch-perfect voices, Kathryn Stockett creates three extraordinary women" (Stockett 2009). The publishers of *The Help* and several African American critics have lauded the authenticity of BVE in the novel. Karen Grisby Bates, a book critic for National Public Radio, wrote that "Stockett masterfully captures both black and white voices with astonishing believability" (Bates 2009). She then goes on to compare it to the novel *To Kill a Mockingbird*. Clearly this African American critic did not find the BVE in the novel offensive. Stockett herself, however, has said that any discussion of the authenticity of her attempt at BVE makes her intensely uncomfortable. She claims that she did not think about how others would perceive the dialogue because she never expected anyone to read her novel (Norris 2009). It is irrelevant whether this claim is true or whether it is retrospective false modesty. The fact remains that the book is published, and hosts of readers have devoured its pages. Could the dialect be accurate if Stockett put so little thought into it?

Cacophonies

The naysayers claim that rather than a novel of heteroglossia and polyphony, *The Help* is an affected clashing of stereotypes. The Association of Black Women Historians issued an open letter condemning the novel. They write that Stockett's attempt at dialect is "an irreverent depiction of black vernacular" (Jones et al. 2011), and they attack a line from the trailer of the film adaptation where Aibileen says, "You is smart, you is kind, you is important." They compare the depiction of black women as akin to the deeply imbedded racism in *Gone with the Wind*. According to them, language played a key role in the stereotyping in both cases (Jones et al. 2011). Furthermore, bell hooks, the outspoken black literary theorist, has lambasted the language of *The Help* in several public forums. If, as the critics claim, *The Help* has

only a caricatured and fractured depiction of the language of black women, how can it accurately portray their deeper complexities and dire struggles amidst that dark hour of American history? If the language is grossly simplified, so also must be the thought processes and, in short, the lives of these women. This would be the literary equivalent of Vaudeville's blackface.

Methodology

Since my research is grounded in a textual analysis, there are no living subjects that need to be taken into account. Instead, I will analyze the first two chapters about Aibileen (approximately 11,000 words) in the novel by evaluating Stockett's attempts at BVE. My judgment will be largely qualitative, but will be rooted in the grammar written by Lisa Green in her book *African American English: A Linguistic Introduction*. I will also provide an approximate percentage of the accuracy of BVE in Stockett's prose.

Verbs

Evaluating the BVE in *The Help* would be too large a task for this paper if I were to examine each usage of every facet of the dialect. It is beyond the scope of this paper to piece together Stockett's intended phonetics based on her orthographic decisions, to look into how frequently certain lexical items were used in Jackson in the 1960s, or to catalogue every nonstandard morphological addition to every noun. Verbs are the heart of language, so this study is rooted in an examination of the verbs in *The Help*.

The verbs of BVE are particularly useful because of their marked difference from standard American English. BVE regularly conjugates the third person singular present form of a verb the same as the other persons in the present tense (e.g., I *go*, you *go*, he *go*). Also, BVE encodes much more information about the aspect of a verb than Standard English does. The aspectual particle *be* for example communicates

habitual or regular action ("I be going to the store on Thursday" means "I usually go to the store on Thursday"). The aspectual participle *done* implies action that is already past ("I done paid the rent" means "I have already paid the rent") (Green 2002). Context should make clear whether Stockett understands these nuances or whether she uses them incorrectly.

Percentage of Accuracy

This is a tricky area to quantify, as the final figure represents more of a general assessment than a definitive commendation or condemnation of Stockett's usage. I evaluated every instance of Stockett's usage of verbs that depart from Standard English and enter into BVE or something attempting to approximate it. I tallied each time she used a nuance correctly as well as each time she forced a structure or form onto the dialect that does not exist in the natural dialect. The final figure represents the percentage of accuracy Aibileen shows. Thus a 0 percent would mean that Aibileen speaks without any authenticity in regard to her BVE verbs. A score of 100 percent would mean that Aibileen speaks with fully authentic BVE verbs.

Threshold of Significance

In all, I observed 2,412 verbs. I divided them into three categories: (1) verbs that are indistinguishable from standard English, (2) verbs that conform to BVE as documented and published by professional linguists, and (3) verbs that are either misuses of BVE or bear no resemblance to either of the former categories. Because of the large sample and the small number of categories, my findings are extremely statistically significant with a p-value of less than 0.0001.

In deciding whether Stockett uses sufficiently accurate BVE in her depiction of Aibileen, I set the bar at 75 percent accuracy. That is, when just considering categories 2 and 3, category 2 must constitute at least 75 percent of the uses. Such a high percentage reflects that

Stockett's accuracy is not simply by chance. Rather, the level of accuracy would ensure that Stockett was at least informed concerning the dialect, if somewhat off on the specifics.

Analysis

To my genuine surprise, Stockett scored quite high in my study. The following table shows the breakdown.

Verb Classification	Total	Percent of Total	Total Accuracy
1. Standard English	1,776	73.63%	—
2. Accurate BVE	592	24.54%	93.1%
3. Misused/False Forms	44	1.82%	6.9%

Table 1. The total occurrences of verbs in each of the three categories, the percent of those occurrences from the total of 2,412, and the percent of accurate BVE

Standard English

Most of Aibileen's verbs align with Standard English. This result should not be surprising since in BVE most verb forms are identical to those in Standard English. Also, not all speakers of BVE use the forms specific to the dialect all the time. The dialect involves large amounts of code switching. It would make sense for Aibileen to use many verbs that align with Standard English.

Accurate BVE

The majority of instances of accurate BVE come from the following categories: copular deletion[2]; unconjugated, third-person singular, present verbs[3]; negations with *ain't*[4]; and the omission of helping verbs.[5] Stockett consistently and accurately uses these forms throughout Aibileen's prose.

Misused/False Forms

Stockett clearly does not grasp, however, the correct use of verbal markers in BVE. She particularly does not use *be* and *done* correctly. Lisa Green commented on the common misunderstanding of *be* among nonnative speakers of BVE. She writes, "Ironically, this *be* is often used incorrectly by the same people who try to show that what is taken as AAE [African American English] is illogical speech" (Green 2002). Although I doubt Stockett is trying to make a case against the logical merits of BVE, she exhibits the lack of understanding Green describes. For example, Aibileen says, "Right next to Belhaven be the downtown and the state capital" (Stockett 2009). The use of *be* here should mean that the circumstance or action is a recurring or regular one. But the sentence "The downtown and state capital are usually right next to Belhaven" is nonsense. The state capital does not habitually locate itself right next to Belhaven.

Stockett also fails to accurately use the verbal marker *done*. For example, in the first paragraph of the first page of the book, Aibileen says, "I done raised seventeen kids in my lifetime." Stockett appears to be using *done* as some kind of emphatic particle or *do*-support, similar to "I *do* look after seventeen kids." But that is not how *done* functions in BVE. The verbal marker indicates action that has already been accomplished. So Aibileen's sentence would most accurately be translated as "I have already raised seventeen kids in my lifetime." However, within the context of the story, I doubt that this was Stockett's intended meaning.

Conclusion

Although critics may attack *The Help* on historical, cultural, ethical, or literary grounds, the novel certainly proves itself linguistically. These findings were contrary to my expectation. When I first casually read through the novel the errors seemed to leap off the page, but once the verbs were all quantified, I must admit that Stockett's accuracy is admirable. Although 93-percent accuracy is not "pitch-perfect," it is, nonetheless, impressive for a nonnative speaker.

The Case for Kathryn Stockett

Stockett exhibits a fairly approximate representation of the dialect. Perhaps then, critics should not be as dismissive of the other content in the book. While Stockett does not write with a perfect voice, she at least writes with an informed voice.

This study may also redeem the character of Aibileen, who comes dangerously close to being the trope that African American filmmaker Spike Lee calls the "magical, mystical negro" (Gonzalez 2001). She has the power to heal through supernatural means, and at one point it is implied that she caused a long-distance, three-month, magical vaginal infection (Stockett 2009, 24). Aibileen further conforms to the trope by serving as the guide to the white protagonist. But clichés aside, Aibileen's language is accurate, and therefore, possibly exonerates her from the label of "magical, mystical negro." On the other hand, just because Stockett depicts one facet of a character accurately, it does not mean that the character is accurate or believable as a whole.

Lingering Issues

There are still some facts that tarnish Kathryn Stockett's novel. Primarily, the book is about a white woman who collaborates with black women to share their stories. The white protagonist, Skeeter, shares the money she makes with her black friends who gave her the source material for the book. As mentioned in the introduction, Stockett clearly used the likeness and history of her brother's maid, Ablene, to write *The Help*, yet unlike Stockett's heroic Skeeter, Stockett has made no effort to share any portion of her substantial earnings with Ablene Cooper.

Limitations and Shortcomings of the Present Work

This study has several limitations. For one, I only analyzed verbs. As useful as verbs are, they do not represent the dialect in its entirety. Stockett forms impossible plurals on several occasions, but these are not factored into my data.

Additionally, this study only analyzes the first two chapters of the book; it is possible that Stockett's usage changes as the book progresses. Also, in a very small minority of instances, I was unable to find sources on the legitimacy of some of Stockett's verb uses. Rather than counting them as either correct or fabricated, I simply left them out of the data. These mystery verbs, however, would only add about a dozen new entries to any of the three categories. Even if these verbs were all "Misused/False Forms," their addition would not significantly change my findings.

Future Work

In the future, more studies could be done to test the accuracy of other aspects of Aibileen's morphology. I did not study adverbs, adjectives, nouns, or pronouns, and I did not comment on any of Aibileen's syntactic structures. There is still plenty of research to be done in each of these areas. Also, I only studied Aibileen. A more fruitful study might include a comparison between Minny and Aibileen. If Aibileen's BVE is more caricatured than Minny's, there could be enough evidence to argue that Aibileen is, in fact, yet another addition to the pantheon of the "mystical, magical negro" trope.

A similar methodology to mine could be applied to the works of the great Southern luminaries such as William Faulkner, Eudora Welty, Flannery O'Connor, or Mark Twain. Such studies could foster an engaging discussion among literary theorists and critics as to whether linguistic accuracy is part of the criteria of great literature, or if we should make concessions for authors in the name of poetic license. At the heart of this debate is the larger question of how far divorced literature can be from language.

References

Bates, Karen Grisby. 2009. "A Nuanced Novel of Race in the Deep South." *NPR Books*. <http://www.npr.org/2011/07/29/105920680/a-nuanced-novel-of-race-in-the-deep-south>

Blakely, Kiri. 2011. "The Help: Maid Should Cash In Instead of Sue." *Forbes*.<http://www.forbes.com/sites/kiriblakeley/2011/08/11/the-help-maid-should-cash-in-instead-of-sue/>

Gonzalez, Susan. 2001. "Director Spike Lee Slams 'Same Old' Black Stereotypes in Today's Films." *Yale Bulletin & Calendar*. <http://www.yale.edu/opa/arc-ybc/v29.n21/story3.html>

Green, Lisa. 2002. *African American English: A Linguisitic Introduction*. New York: Cambridge UP.

"Black Stereotypes in Today's Films." <http://www.yale.edu/opa/arc-ybc/v29.n21/story3.html>

Jones, Ida et al. 2011. "An Open Statement to Fans of The Help." *ABWH*. <http://www.abwh.org/index.php?option=com_content&view=article&id=2%3Aopen-statement-the-help>

Norris, Michelle. 2009. "'The Help' Author Says Criticism Makes Her 'Cringe.'" *NPR*. <http://www.npr.org/templates/story/story.php?storyId=120966815>

Stockett, Kathryn. 2009. *The Help*. New York: Putnam.

Endnotes

1. There are a multitude of names for this dialect, as Lisa Green points out on the sixth page of her work on grammar *African American English* (2002). I choose "Black Vernacular English" rather than "African American English" simply because I think it is a more accurate name. This paper is not concerned with how African immigrants in America speak. I am concerned with how a racial dialect is depicted. Green chooses "African American English" because of the influence of African languages on the dialect, but that is not the concern of the present study.
2. "We [ø] one big ant hill . . . " (Stockett 2009).
3. "Baby Girl hug on my legs all afternoon . . . " (ibid).
4. "I walk in the house, glad I ain't got two broken legs" (ibid).
5. "All a sudden she [ø] talking to me like I'm her best friend" (ibid).

Students

Understanding Faith as an Ontological Metaphor

The concept of faith is abstract. This article discusses how people use metaphors to understand faith. To demonstrate this, the author presents examples of faith metaphors from the Book of Mormon. Using this evidence, the author argues that people view faith in concrete terms in order to comprehend it.

Rachel Clawson

Introduction

The word of God is systematic; there is order and consistency to the gospel of Jesus Christ, for God is "the same yesterday, today, and forever" (1 Nephi 10:18). This consistency is manifest in all aspects of the gospel, from the foundational structure of the Church to the linguistic intricacies of the scriptures, and the usage of the word *faith* in the Book of Mormon is no exception. In the Book of Mormon, faith is discussed in terms of the ontological metaphor map FAITH IS AN ENTITY.* This map extends to the metaphor FAITH IS A POSSESSION, which can be broken down into three additional subcategories: FAITH IS A CONDUIT, FAITH IS GROWABLE/MEASURABLE, and FAITH IS A VEHICLE.

The foundation for this research is based on George Lakoff and Mark Johnson's *Metaphors We Live By*, as well as on a subsequent article by Lakoff entitled "The Contemporary Theory of Metaphor." Examples of the use of the word *faith* are limited to its uses in the Book of Mormon as listed in *A Complete Concordance of the Book of Mormon*, compiled by George Reynolds.

Ontological Metaphors

In a linguistic sense, metaphors go far beyond the typical poetic definition of a literary device that compares two unrelated things. Metaphors are the basis of our conceptual knowledge because they provide a foundation for understanding abstract concepts. Metaphors subtly infiltrate our language, as they are "mostly unconscious, automatic, and used with no noticeable effort" (Lakoff 1992, 245). There are several categories of metaphors. The metaphors of faith in the Book of Mormon are ontological metaphors.

Ontological metaphors are "ways of viewing events, activities, emotions, ideas, etc., as entities and substances" (Lakoff and Johnson

* This paper will follow the standard linguistic practice of marking metaphors by writing them in small capital letters.

1980, 25). In short, they help us understand abstract concepts. There are five specific things that ontological metaphors allow us to do with abstract concepts: refer to, quantify, identify aspects of, identify causes of, and set goals and motivate actions regarding them (26).

INFLATION IS AN ENTITY is an ontological metaphor in which inflation (an abstract concept) is referred to as an entity (a concrete object) that can be physically handled, manipulated, and referenced. This metaphor is used in the phrases, "If there's much *more inflation*, we'll never survive," and "*inflation makes me sick*" (Lakoff and Johnson 1980, 26). At first, these examples may be hard to accept as metaphors. It seems reasonable to conclude that inflation actually increases and decreases and that no metaphor is needed to describe that process; however, the ontological metaphor that is part of the concept of inflation allows us to talk about this inflational growth. We cannot physically touch inflation. We cannot go to the "inflation tree" and measure its height with a yardstick. The ontological metaphor INFLATION IS AN ENTITY is necessary if we are to refer to and quantify this abstract idea.

Ontological metaphors are used in everyday speech more often than we recognize. "Ontological metaphors . . . are so natural and so pervasive in our thought that they are usually taken as self-evident, direct descriptions of mental phenomena. The fact that they are metaphorical never occurs to most of us" (Lakoff and Johnson 1980, 18). When we talk about abstract concepts, we most likely use an ontological metaphor as the foundation for our conversation. Lakoff explains, "As soon as one gets away from concrete physical experience and starts talking about abstractions or emotions, metaphorical understanding is the norm" (1992, 205). Ontological metaphors are necessary for us to explain, teach, and understand abstract concepts like faith.

Metaphorical Mapping

A metaphorical map is a tightly structured schema that underlies a metaphor. A map is made up of two parts: the target domain and the

source domain. The target domain is the abstract concept we are trying to understand, and the source domain is a concrete idea or entity, the source for understanding the target domain. A metaphorical map is represented as TARGET DOMAIN IS SOURCE DOMAIN. For example, in the metaphorical map INFLATION IS ENTITY, *inflation* is the target domain, and the notion of an *entity* is the source domain. This system allows other metaphors to emerge as subcategories of this map.

The metaphorical map LOVE IS A JOURNEY leads to the submetaphors LOVERS ARE TRAVELERS (e.g., "we are going our separate ways") and RELATIONSHIPS ARE VEHICLES (e.g., "this relationship isn't going anywhere"). The metaphorical map and its subcategories are represented as

LOVE IS A JOURNEY
    LOVERS ARE TRAVELERS
    RELATIONSHIPS ARE VEHICLES

Metaphorical maps form the foundation for our knowledge of an abstract idea, and then lead us to subsequent metaphors that guide our thinking (Lakoff 1992).

Faith

Faith is abstract. It cannot be touched or handled. *The Oxford English Dictionary* describes faith as "belief, trust, confidence" (2010). In *A Complete Concordance of the Book of Mormon*, the entry on faith is divided into fourteen segments, with each segment representing a common phrase that includes the word *faith*. These faith phrases follow:

> by faith, exceeding faith, great faith, have faith, in faith, in the faith, our faith, because of their faith, according to their faith, their faith, through faith, thy faith, your faith, faith (Reynolds 1990, 218–219)

With these phrases as a guide, faith can be understood through the ontological metaphor map FAITH IS AN ENTITY. The map then leads us to the submetaphor FAITH IS A POSSESSION, which then breaks down

into three additional metaphors; FAITH IS A CONDUIT, FAITH IS GROWABLE/MEASURABLE, and FAITH IS A VEHICLE. This breakdown of the ontological metaphor of faith can be represented by the following:

FAITH IS AN ENTITY
  FAITH IS A POSSESSION
    FAITH IS A CONDUIT
    FAITH IS GROWABLE/MEASURABLE
    FAITH IS A VEHICLE

All instances of *faith* in the Book of Mormon fit into this metaphorical map.

FAITH IS AN ENTITY

As an underlying metaphorical map, the ontological metaphor FAITH IS AN ENTITY allows us to refer to faith, quantify faith, identify aspects of faith, identify causes of faith, and set goals and motivate actions regarding faith. Through this metaphor, the target domain *faith* is associated with the source domain *entity*, and we can talk of faith as something that we physically touch, handle, and manipulate.

FAITH IS A POSSESSION

In the faith phrases outlined in *A Complete Concordance of the Book of Mormon*, the entity of faith is often referred to as a possession: *have faith*, *our faith*, *their faith*, and *your faith*. These phrases are seen in the following examples:

> If ye *have faith* ye can do all things which are expedient unto me. (Moroni 10:23; italics added)
>
> Because of *thy faith* in Christ . . . (Enos 1:8; italics added)
>
> Give us strength according to *our faith* which is in Christ . . . (Alma 14:26; italics added)

The reference to faith as a possession allows us to refer to faith as a concrete entity that people own, just as they own physical money (e.g., "if you have enough money, you can buy this candy bar"). This suggests that those who own faith can literally touch and hold it.

FAITH IS A CONDUIT

Faith as a possession can be used as a conduit between heaven and earth. Heaven can provide divine assistance to possessors of faith—faith is the channel by which heaven helps man. This is seen in the following faith phrases: *by faith, according to their faith, because of their faith,* and *through faith*. These phrases are used in the following verses:

> All they who wrought miracles wrought them *by faith* . . . (Ether 12:16; italics added)
>
> *Because of their faith* in the Lamb of God their garments are made white in his blood. (1 Nephi 12:10; italics added)
>
> Your hearts are changed *through faith* on his name . . . (Mosiah 5:7; italics added)

In the Book of Mormon, those who possessed faith were given assistance to perform miracles, be cleansed of sin, and have their hearts changed. Divine power reaches owners of faith through and because of their possession.

Faith can also be used for individuals to reach heaven. Faith is the means by which human beings on earth ask for divine help. The phrase *in faith* almost always describes owners of faith soliciting heaven's aid:

> I cry unto my God *in faith,* and I know that he will hear my cry. (2 Nephi 33:3; italics added)
>
> Whatsoever thing ye shall ask *in faith* . . . (Enos 1:15; italics added)
>
> Call on his name *in faith* . . . (Alma 22:16; italics added)
>
> Prayed *in faith*. (Alma 31:38; italics added)

Owners of the entity faith can use their faith as a conduit to reach heaven. This medium is what makes prayers successful; when faith is owned, prayers can be answered.

FAITH IS GROWABLE/MEASURABLE

The metaphor FAITH IS GROWABLE/MEASURABLE is probably the most familiar of the five. The Latter-day Saint primary song "Faith" includes the lyrics "Faith is like a little seed: if planted it will grow" (Jackson 1989). This comparison comes from the metaphor FAITH IS GROWABLE/MEASURABLE.

There are two faith phrases that indicate the presence of this metaphor in the Book of Mormon: *great faith* and *exceeding faith*.

Great faith indicates that faith can be measured:

> So *great faith* have I never seen among all the Jews . . . (3 Nephi 19:35; italics added)

Faith phrases that use *exceeding faith* suggest that there is a perfect faith and that if that faith continues to grow, it can exceed that point of perfection:

> Pray unto him with *exceeding faith* . . . (Jacob 3:1; italics added)

There are other words that are commonly used with *faith* to indicate its ability to grow, to be measured, to reach a perfect state, and then to exceed that perfect state. These words include *exercise*, *perfect*, *increase*, and *dormant*:

> They *exercise faith* in him . . . (1 Nephi 7:12; italics added)
>
> Having *perfect faith* . . . (2 Nephi 9:23; italics added)
>
> Would not this *increase your faith*? I say unto you, yea: nevertheless *it hath not grown up to a perfect knowledge*. (Alma 32:29; italics added)
>
> Your *faith is dormant* . . . (Alma 32:34; italics added)

When the metaphor FAITH IS GROWABLE/MEASURABLE is described in the Book of Mormon, it always falls into the hands of the owner of the faith to make it grow. This responsibility is never placed on anyone else.

FAITH IS A VEHICLE

The phrase *in the faith* is usually surrounded by the words *continue*, *steadfast*, *firm*, *endure*, and *end*. Literally, *in the faith* suggests someone is standing inside the entity *faith*—FAITH IS A VEHICLE. The owner of faith can enter into their faith, be encompassed by it, and be aided by it in continuing to the end of the journey:

> Continue *in the faith* even unto the end of his life . . . (Mosiah 4:6; italics added)
>
> Standing steadfastly *in the faith* . . . (Mosiah 4:11; italics added)
>
> And they were firm *in the faith* of Christ, even unto the end. (Alma 27:27; italics added)
>
> Stand fast *in the faith* . . . (Alma 45:17; italics added)

The literal notion "*in* the faith" also suggests that when the owner of faith is standing in this vehicle, they are protected. Their faith surrounds them and shields them from the bumps and bruises that they receive on this earth; faith makes it possible for them to endure to the end.

Conclusion

Faith as represented in the Book of Mormon is an ontological metaphor. All references to faith are based on the ontological metaphor map FAITH IS AN ENTITY. Without this underlying system in place, it would be impossible for us to discuss faith at all. We would be unable to refer to it or quantify it. With this map as a foundation, we can understand faith as a possession that serves as a conduit between man and heaven, a possession that grows and is measurable, and a possession that serves as a vehicle for enduring to the end.

The ontological metaphor of faith is what allows us to understand this abstract idea and apply the saving principle to our lives. Through these metaphors, we learn how to use, develop, and be protected by our faith. It is our faith that will bring us success in enduring to the end.

References

Oxford English Dictionary, s.v. "Faith." 2010. London: Oxford University Press.

Jackson, Beatrice G. 1989. "Faith." *Children's Songbook.* Salt Lake City: The Church of Jesus Christ of Latter-day Saints.

Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By.* Chicago: University of Chicago.

Lakoff, George. 1992. "The Contemporary Theory of Metaphor." *Metaphor and Thought,* ed. Andrew Ortony, 202–251. Cambridge: Cambridge University Press.

Reynolds, George. 1990. *A Complete Concordance of the Book of Mormon.* Salt Lake City: Kessinger Publishing.

The Nonnative Speaker's Challenge: Multiword Verbs

This article explores the challenges that multiword verbs present to nonnative speakers of English. The author explains why these linguistic constructions present problems to nonnative English speakers and how English teachers could address these challenges in the classroom.

Katie Cutler

Typically in English, word classifications are thought to apply only to single words. Sometimes, however, several words function as a single unit, fulfilling a single purpose. One example of this construction is multiword verbs. Multiword verbs consist of a verb combined with either an adverb or a prepositional particle, which creates a verb with a different syntactic or semantic meaning from the meanings of the two words in isolation. The idiomatic character of these constructions challenges nonnative English speakers, especially speakers whose primary language does not contain comparable constructions. This paper will define multiword verbs, address the challenges encountered by nonnative speakers when faced with these constructions, and give approaches to teaching nonnatives that could aid in the mastery of these idiomatic expressions.

Multiword verbs, which are made up of a verb and a prepositional or adverbial particle, are used frequently in both standard and nonstandard English. Linguists agree that multiword verbs tend to occur more often in the informal spoken register and are colloquial in tone (Siyanova 2007, 119). The lexical verbs that occur in multiword verbs are generally more common and are mostly associated with physical movement, such as *come, fall, give, go, keep, make, put,* and *take* (Quirk 1985, 1155). There are four main types of multiword verbs: intransitive phrasal, transitive phrasal, prepositional, and phrasal-prepositional. Intransitive phrasal verbs do not have an object: for example, *grow up* in the sentence *you need to grow up* (Biber 2002, 125). Transitive phrasal verbs, on the other hand, do have an object; this object can be placed either directly after the prepositional particle or between the verb and the particle. For example, one could say *He picked up the phone* or *He picked the phone up.* Prepositional verbs do not have the same flexibility, as the object can only come after the prepositional particle. *They depend on her* cannot be rephrased as *They depend her on* (Biber 2002, 125). Additionally, because of the idiomatic nature of multiword verbs, they often have one-word counterparts. These one-word correspondents are more common in the academic

register, while the multiword verbs are generally used more frequently in the spoken register (Biber 2002, 127). For example, *go away* is more colloquial than *leave*; the same is true for *show up* and *appear*, *put out* and *extinguish*, *give in* and *surrender*, and *show off* and *boast* (Liao 2004, 222). The complex nature of multiword verbs often confuses nonnative speakers, causing them to avoid these constructions.

The idiomatic quality of multiword verbs is one reason nonnative English speakers misinterpret them. Nonnative speakers tend to decipher multiword verbs literally, piece by piece. However, since the meaning of most multiword verbs is not derived from the individual meanings of the words, it is more difficult for nonnative speakers to know how to use them, and thus they do not use them as frequently (Siyanova 2007, 120). For example, the phrase *to put up with someone* does not mean to put something in a higher place with another person; rather, it means to tolerate someone. The variety of multiword verb structures also contributes to the confusion. It is especially difficult to differentiate between transitive phrasal verbs and prepositional verbs. Nonnative speakers fear putting the object in the wrong place, and therefore avoid using these constructions altogether (Siyanova 2007, 120).

An additional impediment for nonnative speakers is that multiword verbs are often polysemous, meaning each construction has several different definitions (Trebits 2009, 471). The verb *to bring up* can mean to nurture (She *brought up* the children), to mention (They *brought up* a good point), or to carry up (He *brought up* the books from the basement) (Siyanova 2007, 120). Each of these examples is at a different level of idiomaticity; *bring up*, when it means to carry, is not very challenging to interpret. When it means to nurture, however, this verb becomes more figurative and obscure, confusing nonnative speakers (Siyanova 2007, 119).

When their native language does not contain constructions like multiword verbs, individuals can become confused and avoid these constructions. Multiword verbs are Germanic; Romance languages do not have any similar constructions, which means that a nonnative speaker

whose primary language is Romanic has no familiarity with multiword verbs. In a study done by Dagut and Laufer, Hebrew speakers tended to avoid multiword verbs, especially those with more figurative meanings (Liao 2004, 197). In a follow-up study, Hulstijn and Marchena found that Dutch speakers seemed much more comfortable with multiword verbs than Hebrew speakers. Through their research, they concluded that a contributing factor to these different results was likely the fact that Dutch, a Germanic language, has a construction similar to the English multiword verb while Hebrew does not (Liao 2004, 198).

Hulstijn and Marchena's study revealed another interesting aspect—nonnative avoidance of familiar constructions—which needs more research to be confirmed. The Dutch speakers that were studied showed a tendency to avoid English multiword verbs with literal Dutch language counterparts (Liao 2004, 199). The researchers hypothesized that this is due to Dutch speakers being unsure of which meanings transferred from Dutch and which did not. According to Hulstijn and Marchena, these speakers tended to "play it safe" and avoided those multiword verbs altogether (as cited in Liao 2004, 199). In addition to the previously mentioned idiomatic nature of the multiword verb, it has been suggested that language differences and similarities can lead to speaker avoidance.

Many studies have addressed this nonnative speaker avoidance, and there are numerous resources available to help instructors teach multiword verbs and other idiomatic expressions. The available research on the subject routinely discusses why nonnative speakers need to understand idiomatic expressions, such as multiword verbs, as well as ways speakers can be more effectively taught in a classroom setting. First, multiword verbs and other idiomatic expressions are important for nonnative speakers to learn because such knowledge improves their total comprehension, especially in informal writing or conversation. As multiword verbs are most common in conversation and fiction, when a nonnative speaker does not understand the meaning of the verb, a large part of the meaning of the overall sentence is lost (Biber 2002, 127).

It is important for nonnative speakers to learn and utilize idiomatic expressions if they intend to converse like a native speaker. Although it is grammatically correct to use a one-word counterpart to a multiword verb, it is generally not what a native speaker would use in casual conversation and writing; it tends to sound more forced, contrived, and pretentious (Prodromou 2003, 44). "Full and effective communication in a second language presupposes not only a knowledge of how things are said in that language, but also what is said" (Scott 1964, 489). Once nonnative speakers learn what is said in English and fully master idiomatic expressions, they can artfully use them in more creative, spontaneous ways and use the language to a greater extent.

Researchers have made several suggestions for how one should teach multiword verbs. First, researchers have proposed that some prepositional particles may carry the same meaning among various multiword verbs, and that learning these consistent meanings could help nonnative speakers interpret the multiword verb as a whole. For example, Peter Machonis argues that the particle *up* often denotes maximal effect, as in the verbs *clean up*, *wipe up*, and *tear up* (2009, 253). The particle *around* often denotes aimless behavior in the context of prepositional verbs such as *play around, fool around, mess around,* and *wait around* (Quirk 1985, 1163). The usefulness of this concept as a teaching method is certainly debatable. Although it is an interesting observation and would be a viable topic for further research, many researchers argue that because the prepositional particles do not carry the same meaning in every multiword verb (*up* can also mean that a process has been started or completed like in *round up the cattle*, or it can mean a process has gained a higher intensity like in *speed up the engine*), and there are many multiword verbs that fit into no category at all, it would not be a practical tool and would generally confuse nonnative speakers (Neagu 2002, 123).

In his article "Idiomaticity and the Nonnative Speaker," Luke Prodromou suggests that teachers should not save learning idiomatic expressions for advanced language classes. Idioms must consistently

be taught as the vocabulary is learned in order to follow the pattern of native language acquisition as closely as possible (2004, 44). Often, in a classroom environment, if a student asks about the meaning of a multiword verb or other idiomatic expression, teachers tend to quickly address the meaning in the specific context and move on instead of trying to thoroughly explain the meaning and promote long-term learning (Adkins 1968, 151). It is important to help students gain a comprehension of the nature of idiomatic language; teachers should encourage the utilization of context when a student attempts to learn the meaning of multiword verbs. By using context to guess at the meaning of an idiomatic expression, the meaning is reinforced in the students' mind and they will be more likely to remember it in the future (Adkins 1968, 151).

There are also practice exercises, which nonnative students can do in order to master idiomatic expressions, that have demonstrated efficacy in various observational studies (Adkins 1968, 152). These practice exercises were structured in a five-week program, in which the teacher would present a news story each week, and the students would read through the article individually. They would then review the story as a class and discuss any vocabulary they did not understand, focusing on multiword verbs and other idiomatic expressions. They would define these expressions, discuss literal and figurative meanings, and expound on the effect idiomatic terms had on the tone of the article. Later in the week, the students would have conversations with other students and write their own sentences, using various expressions provided by the teacher. Sharing these dialogues or sentences with the entire class reinforced the meanings for the group as a whole (Adkins 1968, 151). The group scores increased 27.3 percent between the pretest and the posttest, demonstrating a significant improvement in the mastery of idiomatic expressions and a fuller comprehension of how idioms are used (Adkins 1968, 152).

Many native speakers are unaware of the complexity of multiword verbs, but these complex grammar constructions create confusion

and discomfort for nonnative speakers trying to master English. The idiomatic nature of the verbs and of their prepositions causes nonnative speakers to avoid these constructions at all cost, making their language sound stilted and unnatural. Much of the flexibility, humor, and creativity that English offers comes from idiomatic expressions like multiword verbs. Because of this, linguists are studying how best to teach idioms to nonnative speakers, and with reasonable success. Instead of simply branding those who struggle with understanding idioms as slow learners, educators must construct specific programs to accommodate the students' needs as learners and speakers of English (Adkins 1968, 152).

References

Adkins, P. G. 1968. "Teaching Idioms and Figures of Speech to Nonnative Speakers of English." *The Modern Language Journal*, 52(3):148–152.

Biber, D., S. Conrad, and G. Leech. 2002. *Longman Student Grammar of Spoken and Written English*. Essex, England: Longman.

Liao, Y., and Y. J. Fukuya. 2004. "Avoidance of Phrasal Verbs: The Case of Chinese Learners of English." *Language Learning*, 54(2):193–226.

Machonis, P. A. 2009. "Compositional Phrasal Verbs with up: Direction, Aspect, Intensity." *Lingvisticae Investigationes*, 32(2):253–264.

Machonis, P. 2008. "Disambiguating phrasal verbs." *Lingvisticae Investigationes*, 31(2):200-212.

Neagu, M. 2007. "English Verb Particles and their Acquisition: A Cognitive Approach." *Revista Espanola de Linguistica Aplicada*, 20: 121–138.

Prodromou, L. 2003. "Idiomaticity and the Nonnative Speaker." *English Today*, 19(2):42–48.

Quirk, R., et al. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Scott, C. T. 1964. "Literature and the ESL Program." *Modern Language Journal*, 48(8):489–493.

Siyanova, A., and N. Schmitt. 2007. "Native and Nonnative use of Multi-Word vs. One-Word Verbs." *International Review of Applied Linguistics in Language Teaching*. Herndon, 45(2):119–139.

Trebits, A. 2009. "The most frequent phrasal verbs in English language EU documents—a corpus-based analysis and its implications." *ScienceDirect*, 37:470–481.

In Other Words:
Understanding the Shakespearean Dialect

The author discusses ways in which inexperienced readers can better understand Shakespeare. She specifically suggests gaining an understanding of Shakespeare's pronoun usage and sentence structure to better understand his intended meaning.

Claire Ford

Poetry is meant to be read aloud; plays are meant to be performed. Speaking the poetry and prose of a Shakespearean text greatly complements a reader's understanding, but what if a vocalized reading still leaves the reader perplexed? By exploring the pronouns and word order within William Shakespeare's texts, this essay searches for ways an inexperienced reader of Shakespeare can find meaning in his timeless words.

Twenty-first-century ears are accustomed to the colloquialisms and grammatical structures that are characteristic of our current generation. Both formal, scholastic language and informal, everyday speech have definite rules for a grammatical framework that governs the form and structure of sentences. As defined by the eighth edition of Martha Kolln and Robert Funk's *Understanding English Grammar*, this framework includes three meanings of grammar: first, the system of rules in our heads; second, the formal description of those rules; and third, the social implications of usage (2009, 5). According to Kolln's first and third elements of grammar, native English speakers grow up accustomed to the English of the current century. They have an intuitive understanding of grammatical rules that is maintained by interaction with other English speakers. This continuous updating of accepted and understood grammar causes many English readers to struggle to understand Shakespearean language. The year 2012 marks the 396th anniversary of Shakespeare's death—much time has passed since common speech resembled Shakespearean language. Thus, most readers today do not have a naturally trained ear for understanding Shakespeare.

Before a reader can find pleasure in a Shakespearean text, he or she must first become aware of the sounds and patterns used throughout Shakespeare's collection of works. A Shakespearean sonnet, for example, gains meaning as the reader recognizes its rhythm and meter. Just as a seasoned sonnet reader "[develops] an instinct for meter," a learner of Shakespeare can acquire an ear for the sentence structure and grammar of a play (Wheeler 1967, 210). One could dissect each line of each play to discover its meaning, but with Shakespeare's plays ranging from 605 to 1,361 lines, a thorough analysis would take hours,

if not days. Instead, a reader must seek to understand the rhetorical devices that Shakespeare sustained throughout his plays.

Learning to understand Shakespearean texts begins with studying the work of previous grammar scholars. N. F. Blake's *A Grammar of Shakespeare's Language* names vocabulary and grammar as the two constant studies of Shakespeare's works. Shakespeare's vocabulary, the first of these studies, has been heavily criticized and edited almost from the time Shakespeare died in 1616. Ben Jonson, a contemporary of the playwright, said that Shakespeare showed a lack of taste and discrimination because of the way he "[mixed] words from a low level of usage with those more appropriate to the grand rhetoric of other passages" (Blake 2002, 8).

However, more recent analyses deem this same scheme as inventive. Shakespeare recognized the difference between upper- and lower-class word choice and alternated between them to show a change in speaker. Contrary to the belief of Jonson, Shakespeare intentionally employed these different levels of language to identify the status of his speaker and the relationship the speaker has with other people involved in the conversation.

This can also happen in modern English. As quoted by Kolln, Paul Roberts gives the following example:

1. Henry brought his mother some flowers.
2. Henry brung his mother some flowers.

Roberts explains that "we associate sentence 1 with educated people and sentence 2 with uneducated people, . . . [but] educated people do not say sentence 1 . . . because it is better than 2. Educated people say it, and that makes it better" (Kolln and Funk 2009, 8). Because of natural assumptions attached to certain words or grammatical structures, readers of the English language associate certain ways of speaking or writing with a certain education level. These associations are exactly what Shakespeare utilized as he wrote his plays.

One device that Shakespeare employed to indicate relationships between speakers is his use of pronouns. For modern readers,

the pronouns *thou*, *thee*, *thy*, *thine*, and *ye* are a common "point of confusion" because *you* is now used for all second-person references (McDonald 2001, 38). Other languages, however, such as German and Spanish, maintain a distinct difference between informal and formal uses of you: *tú* and *usted* in Spanish and *du* and *Sie* in German, for example. Shakespeare deliberately used shifts between *you* and *thou* as signals that the conversation is becoming more formal or informal. As McDonald (2001) points out, Romeo and Juliet use different *you* forms when they first meet at the Capulet party. Juliet properly addresses the stranger, Romeo, as *you*, whereas Romeo reflects his want to move quickly to the familiar when he uses the more informal form: "Thus from my lips, by *thine*, my sin is purg'd" (1.5.107; italics mine).

The issue of *thou* and *you* can also be approached through the lens of foreign-language-learning schemes. Without a foreign-language background, native English speakers would not know, for example, that many languages have separate forms of *you* reflecting either formal or informal conversation. Though English-grammar students would be accustomed to identifying the appositives, conjunctions, or parallel structures of English sentences, second-language learners are often more deliberate in labeling syntactic elements because they must learn these elements through thoughtful study, as opposed to mere social interaction.

The second main area of study defined by Blake, Shakespeare's grammar, has been much more criticized than his vocabulary. Jonson searched Shakespeare's works for poor grammatical usage and took pride in pointing them out to editors. Systematic studies of Shakespeare's grammar did not begin until the nineteenth century, when linguists began to compare Shakespeare's usage to the English of their own time. Comparisons of this sort remain the major theme for writings about Shakespeare (Blake 2002, 8).

The Bedford Companion acknowledges that the greatest number of difficulties in reading Shakespeare comes from the "unfamiliar order" of words used in his sentences (McDonald, 41). From

Shakespeare's time to the present day, linguists and grammarians have continually redefined the standard word order of a sentence. The first edition of Martha Kolln's *Understanding English Grammar* was published in 1982 and has since been revised into seven subsequent editions, the eighth edition currently being studied in university grammar classrooms. Eight different editions of the same book have been published over a span of only eighteen years—language rules are constantly changing. Today, there is an established word order for basic sentences: subject, verb, direct object or complement, and "various modifiers and other subordinate elements distributed where they won't get in the way" (McDonald, 41).

However, McDonald goes on to say, Shakespeare never seems to be intent on forming basic sentences. He, unlike so many writers throughout history, was "in the business of creating stage pictures out of poetry" and was much less concerned about proper grammatical structure than he was about the overall form, fluidity, and picture a poem evoked (41). He saw the image of his poetic vision and created sentences with structures that would add to that vision. For this reason, many of Shakespeare's sentences violate grammar rules that were taught in the sixteenth century as well as those taught today. Is it then feasible to analyze Shakespeare by the ideas of modern-day grammar books? Blake states that "the natural order of events may be disturbed, . . . [and] it is frequent for a modifier to be placed before another word than that to which it rightfully refers" (Blake 2002, 267). Shakespeare also sometimes employed a "change in overall sense as well as in structure" (Blake, 270). For example, in *Hamlet*, Laertes is talking about how his sword will be poisoned when he says:

> I bought an unction of a mountebank,
> So mortal that, but dip a knife in it,
> Where it draws blood not cataplasm so rare,
> Collected from all simples that have virtue
> Under the moon, can save the thing from death
> That is but scratch'd withal
> (4.7.114–19)

Blake points out that this structure, which a reader might call a "problem," is deliberately employed by Shakespeare to show "speakers who are formulating their thoughts and get lost in the grammatical construction" (Blake, 270). In this way, Shakespeare captured the humanity of his speakers. Rather than glorifying them or portraying them as unnaturally well spoken, Shakespeare depicted them as characters whose minds get naturally caught up in emotion and, as a result, mix up their sentence structure. Once the audience becomes aware that the improper grammatical structure is intentional and meant to normalize the characters, the audience can relate more closely to them. McDonald suggests another example of Shakespeare focusing more on fluidity than on proper syntax: ". . . Sense sure you have, / Else could you not have motion" (3.4.71–72). When a reader is not focused on what the words mean, the deliberate rhythm and meter moves them to feel the meaning instead of literally understanding it. "Prepositional phrases or modifying clauses sometimes intrude in positions that strike us as abnormal," and although they violate the basic sentence structure accepted in English, these phrases and clauses create a rhythm that illustrates Shakespeare's intended emotion (McDonald, 41).

A greater knowledge of Shakespeare's pronoun use and sentence structure helps readers to better understand Shakespeare's language without having to pick every line apart. Shakespeare consistently employed methods such as alternating pronoun forms and varying sentence structure in his writing. Recognizing the regularity of these methods in his writing helps the reader realize that they are intentional, giving the reader a greater ability to relate to Shakespeare's characters and understand his works.

References

Blake, N. F. 2002. *A Grammar of Shakespeare's Language*. New York: Palgrave.

Hope, Jonathan. 2003. *Shakespeare's Grammar*. Great Britain: TJ International Ltd. Padstow.

Kolln, Martha and Robert Funk. 2009. *Understanding English Grammar*. 8th Ed. New Jersey: Pearson Education, Inc.

McDonald, Russ. 2001. *The Bedford Companion to Shakespeare: An Introduction with Documents*. 2nd Ed. Stratford, CT: Bedford/St. Martin's.

Wheeler, Charles B. 1967. *The Design of Poetry*. New York: W.W. Norton & Company.

Self-Help Book Conversation: Is It Cooperative and Realistic?

The Cooperative Principle proposed by Paul Grice in Studies in the Way of Words *provides four maxims—Quality, Relation, Manner, and Quantity—that provide guidelines that enable conversation to be meaningful and effective. Here, the author explores these maxims in relation to sample conversations from a self-help book, examining how the maxims improve or degrade conversation. Additionally, the article discusses the possible synthetic nature of self-help book conversations.*

Hillary Mousley

Unspoken expectations in conversation exist in every culture and language. In *Rules for Argumentation in Dialogues,* Frans H. Van Eemeren and Rob Grottendorst outline the concept of social rules in conversation, highlighting the need for cooperation and coordination and suggesting that violation of these rules can undermine "comprehensibility or acceptability of discourse" (1988, 500). The Cooperative Principle and its corresponding maxims proposed by Paul Grice in *Studies in the Way of Words* provide one model to explore the compatibility (and related acceptability and comprehensibility) of conversation. Outlining four maxims—Quality, Relation, Manner, and Quantity—all under the guiding Cooperative Principle, Grice argues that an observation of these principles of communication will yield conversation that is meaningful and effective for both participants (1989, 26–7).

While analyzing discourse models from Patterson, Grenny, McMillan, and Switzler's self-help text *Crucial Conversations: Tools for Talking When Stakes Are High*, I sought to determine whether conversational models are consistent with Grice's popular maxims of cooperative conversation. Finding that strict adherence to Grice's model does not necessarily correspond with the effectiveness of actual conversation in self-help literature, I argue that while his maxims measure efficiency in speech, they do not form a definite framework for determining the cooperative qualities of discourse.

Furthermore, in questioning the validity of these models as believable conversation, I additionally analyzed these conversation models for features of planned and unplanned discourse. Seeking to determine whether one set of qualities predominated, I hypothesized that a generous inclusion of unplanned discourse characteristics would explain why dialogue found in self-help books seems synthetic and inimitable in daily practice.

Exploration of Modeled Dialogue Using Gricean Maxims

I began my analysis by selecting several conversational dialogues, including models conceived to be examples of both poor and effective communication (see appendix). I then analyzed the models in the context of Grice's maxims for cooperative communication while seeking to determine whether a trend existed between conversations labeled "effective" by their author and those conversations' correspondence to the Gricean principles.

My limited analysis of ineffective speech models using Grice's maxims suggests that regular violation of all four maxims is characteristic of self-help literature's ineffective conversation examples. Dialogue regarded as "poor" by self-help book authors consistently disregards Grice's cooperative principle by violating all four of his maxims. Attributes characterized as typical of poor communication, scattered prominently throughout "ineffective" communication examples, include unsupported and untruthful assertions, superfluous information, random interjections, absent transitions, and obscure expressions. These examples can be generally sorted according to the specific maxims they ignore.

Violating the Quality maxim, poor discourse included information that was unsupported by the knowledge of the speaker, which can be noted in the following conversational model:

> **Wendy**: How could you embarrass me like that? I get one boy to like me, and now he'll never talk to me again! I hate you!
>
> **Parent**: That wasn't a boy. That was a future inmate. You're worth more than that. Why are you wasting your time with him?
>
> **Wendy**: You're ruining my life. Leave me alone!

Here, the implication that "that [boy] was a future inmate" highlights a failure to fufill the requirement of the Quality maxim, which is to "not say that for which you lack adequate evidence" (quoted in Johnstone 2008, 234).

Conversation labeled "ineffective" in self-help literature also consistently contained random interjections and unnecessary information in opposition to the Relation maxim, which requires that contributions "be relevant" (quoted in Johnstone 2008, 234). Wendy's seemingly irrelevant exclamation in the previous conversation ("You're ruining my life. Leave me alone!") illustrates this type of violation, as she ignores the question being posed to her by her parent ("Why are you wasting your time with him?"). A similar example of an unrelated tangent can be found in the following conversation:

> **Clerk**: Did everything go all right with the procedure?
>
> **Patient**: Mostly.
>
> **Clerk**: It sounds like you had a problem of some kind. Is that right?
>
> **Patient**: I'll say. It hurt quite a bit. And besides, isn't the doctor, like, uh, way too old?
>
> **Clerk**: Well aren't you the ungrateful one! The kind doctor devotes his whole life to helping people and now that he's a little gray around the edges, you want to send him out to pasture!

Again, this excerpt of "ineffective" conversation highlights a response ("Well aren't you the ungrateful one!") that does not address the inquiry that the patient poses concerning the doctor's age, thus failing to fulfill the maxim of Relation.

Abundant absent transitions and obscurities in expression throughout ineffective conversation models represent a violation of the maxim of Manner. Phrases in the following exchange ("starting to drive me nuts" and "you lay into me") could arguably have alternatives that are more appropriate and clear and that would achieve a similar purpose, especially in the context of the professional situation between a boss and his employee.

> **Brian**: I'd like to talk to you about your leadership style. You micromanage me, and it's starting to drive me nuts.
>
> **Fernando**: What? I ask if you're going to be done on time and you lay into me with . . .

A wealth of unnecessary information is another common characteristic of uncooperative conversation in self-help literature. This tendency clearly overlooks Grice's Quantity maxim, which calls for contributions to be "as informative as required" (Grice 1989, 26). The clerk's response in the excerpt below is one example of such a situation, where information about the doctor's life and age inappropriately dominates the dialogue and ignores cooperative efficiency:

> **Clerk**: Well aren't you the ungrateful one! The kind doctor devotes his whole life to helping people and now that he's a little gray around the edges, you want to send him out to pasture!

Though dialogue termed "ineffective" (and therefore uncooperative) regularly violates the maxims of Quality, Quantity, Relation, and Manner, models termed "effective" conversation in self-help books did not follow such a clear pattern. Ironically, many of the principles purported by the self-help books to promote optimal conversation contradicted Gricean maxims. Conceptually, Grice's maxims appear to be fairly straightforward and follow common sense, but their application is much more complex; in many cases, effective speech flouted Grice's maxims in order to fulfill a greater communicative purpose.

The Relation maxim, for example, was repeatedly violated by the suggestion implied in dialogue models that speakers should anticipate the thoughts and emotions of their addressees. The "cooperative" exchange between a mother and her child illustrates this proposed principle of communication:

> **Wendy**: What, so you can tell me more ways that I'm screwed up? I've finally got one friend who accepts me, and you're trying to chase him away!
>
> **Parent**: So you feel like I don't approve of you, and your friend is one person who does?
>
> **Wendy**: It's not just you. All my friends have lots of boys who like them. Doug's the first guy who's even called me. I don't know—never mind.

Parent: I can see how you'd feel badly when others are getting attention from boys and you aren't. I'd probably feel the same way.

Wendy: Then how could you embarrass me like that?!

Parent: Honey, I'd like to take a stab at something here. I wonder if part of the reason you've started dressing differently and hanging out with different friends is because you're not feeling cared about and valued by boys, by your parents, and by others right now. Is that part of it?

Wendy: (*Sits quietly for a long time*) Why am I so ugly? I really work on how I look but . . .

In this conversation, the parent's attempt to relate to her daughter and understand her situation ("Honey, I'd like to take a stab at something here") marks the use of a tool encouraged by self-help authors and simultaneously forms an unrelated tangent to Wendy's question ("Then how could you embarrass me like that?!"). Likewise, Wendy's response is in the form of a question to the inquiry of her mother ("Why am I so ugly?"), which seems to pose another contradiction to the Relevance maxim.

Yet another modeled technique, which is purported to improve the quality of discourse while contradicting Grice's maxims, is the use of repetition to clarify, sympathize, reinforce, or convey understanding. This practice is modeled in the following excerpt of an "effective" conversational model:

Sister 1: Do you think I'm being unfair? That I'm not acknowledging your contributions?

Sister 2: It's just that I know I wasn't around much in the last couple of years. I've had to travel a lot for work. But I still visited whenever I could, and I sent money every month to help contribute to Mom's care. I offered to help pay to bring in a nurse if you thought it was necessary. I didn't know you felt you had an unfair share of the responsibility, and it seems like your asking for more money is coming out of nowhere.

> **Sister 1**: So you feel like you were doing everything you could to help out and are surprised that I feel like I should be compensated?
>
> **Sister 2**: Well, yes.

In this case, Sister 1's series of questions represent her attempt to establish an understanding of Sister 2's point of view. This dialogue is presented as ideal communication; however, Sister 1's restatement of what she has gleaned from Sister 2's lengthy explanation seemingly contradicts the Quantity maxim, which would call for conciseness in speech.

Despite a lack of compliance to the Relation and Quantity maxims, communication labeled as "effective" in the self-help book did consistently correspond with the Quality and Manner maxims in my analysis. The Quality maxim is signified by the following suggestion from a parent during a modeled exchange, in which the parent hedges his or her assertions ("Honey, I'd like to take a stab at something here"), admitting that he or she is unsure of the claim being made and attempting to provide information known to be truthful:

> **Parent**: Honey, I'd like to take a stab at something here. I wonder if part of the reason you've started dressing differently and hanging out with different friends is because you're not feeling cared about and valued by boys, by your parents, and by others right now. Is that part of it?

Because effective dialogue models are exceptionally organized and clearly articulated, arguably at the expense of establishing any sense of realism, they follow the tenets of the Manner maxim, which calls for brief and orderly speech that avoids obscurity and ambiguity (Grice, 27). According to such adherence to these basic principles, cooperative conversational models also portrayed participants who generally appeared calmer and more emotionally stable.

Grice's Maxims: The Most Effective Model for Analysis?

Though regularly used by discourse analysts, analyses relying heavily on Gricean maxims pose a variety of complications, which were highlighted during my own research experience.

In *Grice's Maxims: "Do the Right Thing,"* Robert E. Frederking argues that Grice's maxims are "hopelessly vague" and "too general" for direct implementation in computational natural language systems (1996, 1). My own analysis highlights dilemmas consistent with these claims, finding Grice's framework to be much more suited for a conceptual understanding, and ill-suited for in-depth analysis. My attempts to comply a discourse with notions of relevance and quantity were complicated by my lack of knowledge of the conversation's context.

As a result, my own judgment of adequate relation and quantity was largely uninformed and undoubtedly impaired. Earlier judgments I made, such as determining the violation of the Quality maxim on the grounds that Wendy's mother did not know enough about the boy her daughter had been seeing to pinpoint him as a "future inmate," could be easily invalidated and disproved. Were there greater contextual information suggesting that the mother does indeed have a thorough understanding of the young man's character, consequently making the conversation in agreement with the Quality maxim, my analysis of this communicative model would shift considerably. Similarly, determining the appropriate quantity of information in dialogues necessary to adhere to Grice's cooperative principle, such as the amount of detail that needed to be disclosed about the mother's health care, was impossible (appendix, example 2).

In addition to the general vagueness of Grice's maxims, which complicates applying them in conversation, their essential reliance on a simple economy model that equates cooperation with efficiency makes them ineffective tools for analyzing the effectiveness of a conversation.

Based upon principles of relevancy and limited quantity, Grice's model in many senses seems to undermine the complexity of discourse, forming guidelines for conversation that are minimal, but not necessarily cooperative. The conversational model between Wendy and her parent (appendix, example 1) illustrates such a tendency, as it presents an example of a mother technically violating the Relation maxim in an attempt to develop a common understanding with her daughter. Likewise, this pattern can be observed in the following selection from the same dialogue model:

> **Wendy**: What, so you can tell me more ways that I'm screwed up? I've finally got one friend who accepts me, and you're trying to chase him away!
>
> **Parent**: So you feel like I don't approve of you, and your friend is one person who does?

The emotionless exchange between sisters (appendix, example 2) illustrates yet another example of dialogue that, though technically acting in accordance with Grice's maxims of Quality and Relation in its strict efficiency, obviously lacks the emotional and social complexity that would normally characterize a discussion concerning the living arrangements and care of an elderly parent. Therefore, even though such a dialogue might be termed cooperative by Grice, its complete ignorance of the emotional elements of discourse would lessen its effectiveness if such a conversation were actually attempted.

Believability and the Features of Planned and Unplanned Discourse

The art of creating a believable and naturally flowing dialogue was a process I also explored in my analysis of self-help book dialogue. Comparing ineffective and effective speech models with the features

of planned and unplanned discourse outlined by Ochs, Tannen, and Chafe (quoted in Johnstone 2008, 213), I sought to determine whether the presence or absence of these features corresponds to the believability of a dialogue as naturally-occurring, spontaneous speech.

Several attributes representative of relatively unplanned discourse—namely, a reliance on immediate context ("How could you embarrass me like that?"), preference of deictic modifiers ("So it's *these* expenses you're worried about covering?"), and greater coordination ("You did a lot to help out, *and* I realize that it was expensive to visit as often as you did.")—were consistently present in both ineffective and effective self-help dialogue examples.

Even more prominent in speech models, however, were features of constructed discourse. Predictably, all the provided dialogue displayed a scarcity of repair mechanisms—a feature expected for any published discourse. Additionally, the speech displayed the use of relative clauses ("boys who like them"), the generous use of tenses beyond the present ("cared," "valued," "offered," "felt"), more attributive adjectives ("kind doctor," "future inmate"), less repetition and parallelism, and more compact, dense, integrated discourse.

Even though both the ineffective and the effective dialogue included features common to constructed dialogue, there was only one clear difference between them in morphological and syntactical complexity. "Effective" communication displayed more intricate, developed, and mature syntax ("I opted not to pay for professional home health care because Mom was more comfortable with me taking care of her, and I didn't mind that."), particularly when juxtaposed with the markedly juvenile syntax of ineffective models ("That wasn't a boy. That was a future inmate."). While ineffective communication was formed by simple structures combining a subject, verb, and complement, effective and sophisticated communication models incorporated use of subordination, coordination, prepositional phrases, and direct objects.

The fact that a number of planned features were found to occur in modeled dialogues offers a possible explanation for why self-help

book conversations may appear so scripted and unnatural, and why they are difficult and awkward to apply in practice. Additionally, a characteristic that is unaccounted for by these criteria but also offers a potential explanation for dialogue feeling artificial, is the general lack of interruptions. This creates a strict turn-taking style of speech that is rarely found in everyday conversation, as is exemplified in the effective conversation model between sisters (appendix, example 2).

The tendency of self-help book dialogues to condemn emotional reactions, explanations, or outbursts in dialogue termed effective may likewise contribute to the almost foreign-feeling nature of scripted self-help dialogue. "Effective" dialogue void of any exclamative punctuation, which typically indicates some level of emotional involvedness in conversation, is heavily dominated by declarative statements and occasional interrogations. In comparison to "ineffective" dialogue models, which include several exclamations, self-help book authors seem to inadvertently imply that emotional outbursts or interjections are characteristic of poor communication. By doing so, they may actually decrease the usefulness of conversational models to real life.

Conclusion

Though they are a fascinating conceptual framework, Grice's Cooperative Principle and corresponding maxims are, at best, weakly capable of determining the effectiveness of conversation, particularly in analysis of self-help book dialogue. Since they provide a bare economy model, Grice's maxims fail to take into account the emotional and social complexities of conversation, making their strict application to speech unrealistic. Furthermore, the display of several planned speech qualities identified by Ochs, Tannen, and Chafe may explain why dialogue models, particularly those termed "effective," feel unnatural and lack a sense of the natural spontaneity of conversation.

References

Frederking, Robert E. 1996. *Grice's Maxims: "Do the Right Thing."* Presented at the Computational Implicature workshop at the AAAI-96 Spring Symposium Series, Stanford. Print.

Grice, Paul H. 1989. *Studies in the Way of Words*. Cambridge: Harvard University Press.

Johnstone, Barbara. 2008. *Discourse Analysis.* 2nd ed. Malden: Blackwell Publishing.

Patterson, Kerry, et al. 2002. *Crucial Conversations: Tools for Talking When Stakes Are High.* New York: McGraw-Hill. Print.

Van Eemeren, Frans H. and Rob Grootendorst. 1988. "Rules for Argumentation in Dialogues." *Argumentation*. 2(4): 499–510.

Appendix

Effective Conversation Models Extracted from *Crucial Conversations*

Example 1 (pp. 154–55)

Parent: (*Tapping on door*) Wendy? May I talk with you please?

Wendy: Whatever.

(*Parent enters her room and sits on her bed.*)

Parent: I'm really sorry for embarrassing you like that. That was a bad way to handle it.

Wendy: It's just that you do that a lot. It's like you want to control everything in my life.

Parent: Can we talk about that?

Wendy: (*Sounding angry*) It's no big deal. You're the parent, right?

Parent: From the way you say that, it sounds like it is a big deal. I really would like to hear what makes you think I'm trying to control your life.

Wendy: What, so you can tell me more ways that I'm screwed up? I've finally got one friend who accepts me, and you're trying to chase him away!

Parent: So you feel like I don't approve of you, and your friend is one person who does?

Wendy: It's not just you. All my friends have lots of boys who like them. Doug's the first guy who's even called me. I don't know—never mind.

Parent: I can see how you'd feel badly when others are getting attention from boys and you aren't. I'd probably feel the same way.

Wendy: Then how could you embarrass me like that?!

Parent: Honey, I'd like to take a stab at something here. I wonder if part of the reason you've started dressing differently and hanging out with different friends is because you're not feeling cared about and valued by boys, by your parents, and by others right now. Is that part of it?

Wendy: (*Sits quietly for a long time*) Why am I so ugly? I really work on how I look but . . .

Example 2 (pp. 190–92)

Sister 1: The way you say that makes it sound like maybe that suggestion isn't okay with you. Is there something I'm missing?

Sister 2: No—if you feel like you deserve more than I do, you're probably right.

Sister 1: Do you think I'm being unfair? That I'm not acknowledging your contributions?

Sister 2: It's just that I know I wasn't around much in the last couple of years. I've had to travel a lot for work. But I still visited whenever I could, and I sent money every month to help contribute to Mom's care. I offered to help pay to bring in a nurse if you thought it was necessary. I didn't know you felt you had an unfair share of the responsibility, and it seems like your asking for more money is coming out of nowhere.

Sister 1: So you feel like you were doing everything you could to help out and are surprised that I feel like I should be compensated?

Sister 2: Well, yes.

Sister 1: You're right. You did a lot to help out, and I realize that it was expensive to visit as often as you did. I opted not to pay for professional home health care because Mom was more comfortable with me taking care of her, and I didn't mind that. On top of that, there were some incidental expenses it doesn't sound like you were aware of. The new medication she was on during the last eighteen months was twice as expensive as the old, and the insurance only covered a percentage of her hospital stays. It adds up.

Sister 2: So it's these expenses you're worried about covering? Could we go over these expenses to decide how to cover them?

Sister 1: I've kept a record of all the expenses that went over the amount that both of us agreed to contribute. Can we sit down tomorrow to go over those and talk about what's fair to reimburse me for?

Sister 2: Okay. We'll talk about the estate and write up a plan for how to divide things up.

Ineffective Conversation Models Extracted from *Crucial Conversations*

Example 3 (pp. 144)

Clerk: Did everything go all right with the procedure?

Patient: Mostly.

Clerk: It sounds like you had a problem of some kind. Is that right?

Patient: I'll say. It hurt quite a bit. And besides, isn't the doctor, like, uh, way too old?

Clerk: Well aren't you the ungrateful one! The kind doctor devotes his whole life to helping people and now that he's a little gray around the edges, you want to send him out to pasture!

Example 4 (p. 127)

Brian: I'd like to talk to you about your leadership style. You micromanage me, and it's starting to drive me nuts.

Fernando: What? I ask if you're going to be done on time and you lay into me with . . .

Example 5 (p. 154)

Wendy: How could you embarrass me like that? I get one boy to like me, and now he'll never talk to me again! I hate you!

Parent: That wasn't a boy. That was a future inmate. You're worth more than that. Why are you wasting your time with him?

Wendy: You're ruining my life. Leave me alone!